

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 71.6116 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-1  
Perfect score: 479  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KKALEKTEADLKKAHPEPE 98

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431	90.0	194	2 AAW14584	Aaw14584 Streptococ
2	431	90.0	194	7 ABW02618	Abw02618 Db16ac pn
3	431	90.0	550	8 ADK48356	Adk48356 Streptococ
4	431	90.0	550	8 ADK48356	Adk48356 Streptococ
5	431	90.0	8991	6 ABW08487	Abw08487 S. pneumo
6	426	88.9	183	2 AAW14570	AAW14570 Streptococ
7	426	88.9	183	7 ABW02604	Abw02604 Bg9739c p
8	424	88.5	168	7 ABW02609	Abw02609 L81905c p
9	408.5	85.3	167	2 AAW14575	AAW14575 Streptococ
10	393	82.0	166	2 AAW14568	AAW14568 Streptococ
11	393	82.0	166	7 ABW02602	Abw02602 Bg8743c p
12	383.5	80.1	185	2 AAW14566	AAW14566 Streptococ
13	383.5	80.1	185	7 ABW02600	Abw02600 Ac94c pne
14	339.5	70.9	204	2 AAW14571	AAW14571 Streptococ
15	339.5	70.9	204	7 ABW02605	Abw02605 Bf1019c p
16	331.5	69.2	170	7 ABW02614	Abw02614 Rct135c p
17	331.5	69.2	181	7 ABW02596	Abw02596 0922134c
18	331.5	69.2	865	6 ABW08489	Abw08489 S. pneumo
19	331.5	69.2	929	2 AAW14593	AAW14593 Streptococ
20	331.5	69.2	929	2 AAW14384	AAW14384 S. pneumo
21	328.5	68.6	188	2 AAW14580	AAW14580 Streptococ
22	328.5	68.6	188	7 ABW02613	Abw02613 Rct129c p
23	323.5	67.5	198	7 ABW02615	Abw02615 Rxc1c pneu
24	323.5	67.5	315	2 AAY04375	AAy04375 Streptococ
25	323.5	67.5	619	2 AAR63437	AAR63437 Pneumococ

## ALIGNMENTS

## RESULT 1

AAW14584  
ID AAW14584 standard; protein; 194 AA.  
XX AC AAW14584;  
XX DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX DE Streptococcus pneumoniae PspA central region.  
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX OS Streptococcus pneumoniae; strain Db16.  
XX FH Key Location/Qualifiers  
FT Misc-difference 61  
FT /note= "unidentified amino acid"  
XX PN WO9709994-A1.  
XX PD 20-MAR-1997.  
XX PD 16-SEP-1996; 96WO-US014819.  
XX PR 15-SEP-1995; 95US-00529055.  
XX PA (UABR-) UAB RES FOUND.  
XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX DR WPI; 1997-202002/18.  
XX DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX Example 6; Fig 13; 296pp; English.

This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain Db16. Comparison of the N-terminal and central regions (AAW14533-57 and AAW14562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX  
SQ Sequence 194 AA;

Query Match 90.0%; Score 431; DB 2; Length 194;  
Best Local Similarity 93.9%; Pred. No. 3.7e-32;  
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
|||  
Db 1 LKEIDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
|||  
Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAVHEP 97  
|||  
Db 61 XSDGEQAQYLAAREEDLIAKAELEQTEADLKKAVNEP 99  
|||

RESULT 2  
ID ABW02618 standard; protein; 194 AA.  
XX  
AC ABW02618;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Db16ac pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.

Key Location/Qualifiers  
FT Misc-difference 1..194 /note="Xaa = Unknown amino acid"  
XX  
XX US6592876-B1.  
XX  
XX 15-JUL-2003.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX  
XX 20-APR-1993; 93US-00048896.  
XX  
XX 06-JUN-1995; 95US-00465746.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.  
XX  
XX Immunological composition for obtaining expression products used for  
XX detecting the presence of Streptococcus pneumoniae or its strain,  
XX comprises at least two different full length isolated gene encoding  
XX pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 64; 121pp; English.

CC The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal  
CC surface protein A (PspA) central region. This sequence is used in the  
CC exemplification of the invention

XX  
SQ Sequence 194 AA;

Query Match 90.0%; Score 431; DB 7; Length 194;  
Best Local Similarity 93.9%; Pred. No. 3.7e-32;  
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
|||  
Db 1 LKEIDSDSDYVKEGFRAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
|||  
Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAVHEP 97  
|||  
Db 61 XSDGEQAQYLAAREEDLIAKAELEQTEADLKKAVNEP 99  
|||

RESULT 3  
ADK48356  
ID ADK48356 standard; protein; 550 AA.  
XX  
AC ADK48356;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Streptococcus pneumoniae protein, Seq ID No 4871.  
XX  
KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
XX  
OS Streptococcus pneumoniae.  
XX  
XX US6699703-B1.  
XX  
XX 02-MAR-2004.  
XX  
XX 26-MAY-2000; 2000US-00583110.  
XX  
XX 02-JUL-1997; 97US-0051553P.  
XX  
XX 12-MAY-1998; 98US-0085131P.  
XX  
XX 30-JUN-1998; 98US-00107433.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;  
XX WPI; 2004-212399/20.  
XX  
XX N-PSDB; ADK45695.  
XX  
XX New nucleic acid molecules and polypeptides useful for diagnosing,  
XX preventing and treating pathological conditions resulting from bacterial  
XX infection, e.g. Streptococcus pneumoniae infection, and in drug  
XX screening.  
XX  
XX Disclosure; SEQ ID NO 4871; 301pp; English.

CC The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.

XX  
SQ Sequence 550 AA;

Query Match 90.0%; Score 431; DB 8; Length 550;  
Best Local Similarity 93.0%; Pred. No. 1.3e-31;  
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 60  
DB 144 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 203  
QY 61 NSDGEQA-QYLAAREEDL-AKKALEKTEADLKKAHVEPE 98  
DB 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 243

RESULT 4  
ID ADR95223 standard; protein; 550 AA.  
XX ADR95223;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE Novel S. pneumoniae protein sequence, SEQ ID 3858.  
XX  
KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;  
KW bacterial infection.  
XX  
OS Streptococcus pneumoniae.  
XX  
PN US6800744-B1.  
XX  
PD 05-OCT-2004.  
XX  
PF 30-JUN-1998; 98US-00107433.  
XX  
PR 02-JUL-1997; 97US-0051553P.  
PR 12-MAY-1998; 98US-0085131P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
DR WPI; 2004-697205/68.  
DR N-PSDB; ADR92620.  
XX  
XX New isolated nucleic acid encoding a Streptococcus pneumoniae  
PT polypeptide, useful for diagnosing, preventing and/or treating  
PT pathological conditions resulting from the bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 3858; 151pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising a sequence  
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its  
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,  
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,  
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,  
CC ADR91886, ADR92197, ADR92334, ADR93039, ADR93079, ADR92366, ADR92650 or  
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide  
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is  
CC hybridisable under high stringency conditions to the nucleotide sequence.  
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.  
CC Also included are a recombinant expression vector comprising the isolated  
CC nucleic acid cited above operably linked to a transcription regulatory  
CC element, a cell comprising the recombinant expression vector and a probe  
CC comprising at least 20 consecutive nucleotides of the nucleotide  
CC sequences as cited above. The methods and compositions of the present  
CC invention are useful for the diagnosis, prevention and/or treatment of  
CC pathological conditions resulting from bacterial infection by  
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and  
CC otitis media. The present sequence is one of the 2603 disclosed S.  
CC pneumoniae protein sequences. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.  
XX  
SQ Sequence 550 AA;

Query Match 90.0%; Score 431; DB 8; Length 550;

Best Local Similarity 93.0%; Pred. No. 1.3e-31;  
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2;  
QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 60  
DB 144 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 203  
QY 61 NSDGEQA-QYLAAREEDL-AKKALEKTEADLKKAHVEPE 98  
DB 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 243

RESULT 5  
ABU08487  
ID ABU08487 standard; protein; 8991 AA.  
XX  
AC ABU08487;  
XX  
DT 24-JUN-2003 (first entry)  
XX  
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
XX  
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
KW antibacterial.  
XX  
OS Streptococcus pneumoniae.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1..8991  
FT /note= "All Xaa residues within this sequence are  
FT unknown"  
XX  
PN US6500613-B1.  
XX  
PD 31-DEC-2002.  
XX  
PF 16-SEP-1996; 96US-00714741.  
XX  
PR 15-SEP-1995; 95US-00529055.  
XX (UVAL-) UNIV ALABAMA.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
XX WPI; 2003-361534/34.  
XX  
PT Isolated PspC amino acid sequence used as polymerase chain reaction or  
PT hybridization probe, comprises pneumococcal surface protein having alpha-  
PT helical, proline rich and repeat regions.  
XX  
PS Disclosure; Col 145-188; 186pp; English.  
XX  
CC The present invention relates to the isolation of Streptococcus  
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
CC like protein having alpha-helical, proline rich and repeat regions. The  
CC PspC and PspA proteins may be used in a vaccine to protect against  
CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
CC PspA may be used for the expression of the proteins, and as PCR primers  
CC or hybridisation probes. The present sequence represents S. pneumoniae  
CC PspA protein  
XX  
SQ Sequence 8991 AA;

Query Match 90.0%; Score 431; DB 6; Length 8991;  
Best Local Similarity 93.9%; Pred. No. 3.5e-30;  
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;  
QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 60  
DB 7537 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKQVEDPK 7596

```
QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEP 97
Db 7597 XSDGEQAQGYLAAAEEDLAKKAELQTEADLKKAVNEP 7635

RESULT 6
AAW14570
ID AAW14570 standard; protein; 183 AA.
XX
AC AAW14570;
XX
XX 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Bg9739.
OS
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 183 AA;

Query Match 88.9%; Score 426; DB 2; Length 183;
Best Local Similarity 92.0%; Pred. No. 1e-31;
Matches 92; Conservative 2; Mismatches 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVEDFK 60
Db 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVEDFK 60

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEP 98
Db 61 NSDGEQAQGYLAAAGEDIKKAELKAEADLKKAVDEPE 100

RESULT 8
ABW02609
ID ABW02609 standard; protein; 168 AA.
XX
XX ABW02609;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
```

ABW02604 standard; protein; 183 AA.

ABW02604;

12-FEB-2004 (first entry)

Bg9739c pneumococcal surface protein A (PspA) central region.

Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine; immunological; gene therapy; immunostimulant.

Unidentified.

US6592876-B1.

15-JUL-2003.

15-SEP-1995; 95US-00529055.

20-APR-1993; 93US-00048896.

08-JUN-1995; 95US-00465746.

(UABR-) UAB RES FOUND.

Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

Example 6; SEQ ID NO 50; 121pp; English.

The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is Bg9739c pneumococcal surface protein A (PspA) central region. This sequence is used in the exemplification of the invention

Seq Sequence 183 AA;

Query Match 88.9%; Score 426; DB 7; Length 183;

Best Local Similarity 92.0%; Pred. No. 1e-31;

Matches 92; Conservative 2; Mismatches 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVEDFK 60

Db 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVEDFK 60

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEP 98

Db 61 NSDGEQAQGYLAAAGEDIKKAELKAEADLKKAVDEPE 100

RESULT 8

ABW02609

ID ABW02609 standard; protein; 168 AA.

XX

XX ABW02609;

AC

XX

XX 12-FEB-2004 (first entry)

DT

XX



DE L81905c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..168  
 FT /note= "Xaa = Unknown amino acid"  
 XX  
 XX  
 XX US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 55; 121pp; English.  
 PS  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies), or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is L81905c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 168 AA;  
 Query Match 88.5%; Score 424; DB 7; Length 168;  
 Best Local Similarity 92.0%; Pred. No. 1.4e-31;  
 Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDQAQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDQAQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 QY 61 NSDGEQA-QYLAABEDL-AKAELEKTEADLKAVHEPE 98  
 DB 61 NSDGEQAQYLAABEDLIAKAEADLKAVDEPE 100  
 RESULT 9  
 AAW14575  
 ID AAW14575 standard; protein; 167 AA.  
 XX  
 AC AAW14575;  
 XX  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain L81905.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 37  
 FT /note= "unidentified amino acid"  
 FT  
 FT Misc-difference 41  
 FT /note= "unidentified amino acid"  
 FT  
 FT Misc-difference 83  
 FT /note= "unidentified amino acid"  
 XX  
 XX WO9709994-A1.  
 PN  
 XX 20-MAR-1997.  
 PD  
 XX 16-SEP-1996; 96WO-US014819.  
 PF  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 PI  
 XX WPI; 1997-202002/18.  
 DR  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 PT  
 XX Example 6; Fig 13; 296pp; English.  
 PS  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain AAL81905.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 167 AA;  
 Query Match 85.3%; Score 408.5; DB 2; Length 167;  
 Best Local Similarity 91.0%; Pred. No. 3.8e-30;  
 Matches 91; Conservative 0; Mismatches 6; Indels 3; Gaps 3;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDQAQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDQAQKLS-LEEXSDKXDELDAEIAKLEKDVDFK 59  
 QY 61 NSDGEQA-QYLAABEDL-AKAELEKTEADLKAVHEPE 98  
 DB 60 NSDGEQAQYLAABEDLIAKAEADLKAVDEPE 99  
 RESULT 10  
 AAW14568  
 ID AAW14568 standard; protein; 166 AA.  
 XX  
 AC AAW14568;  
 XX  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
OS Streptococcus pneumoniae; strain Bg8743.  
XX  
PN WO9709994-A1.  
XX  
PD 20-MAR-1997.  
XX  
XX 16-SEP-1996; 96WO-US014819.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
XX WPI; 1997-202002/18.  
XX  
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX  
XX Example 6; Fig 13; 296pp; English.  
XX  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 166 AA;  
Query Match 82.0%; Score 393; DB 2; Length 166;  
Best Local Similarity 85.0%; Pred. No. 1.1e-28;  
Matches 85; Conservative 6; Mismatches 7; Indels 2; Gaps 2;  
QY 1 LKEIDESSEDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVDFFK 60  
DB 1 LKEIDESSEDYVKEGLRAPLQSKLDKAKKLSKLELSKDIDELDAEIAKLEKVDGDFP 60  
QY 61 NSDGEQA-QYLAAAEEDL-AKKALEKTEADLKAVHEPE 98  
DB 61 NSDGEQAQYLVAAEKDLDAKEAEELGNTGADLKAVDEPE 100  
RESULT 11  
ABW02602  
ID ABW02602 standard; protein; 166 AA.  
XX  
XX ABW02602;  
XX  
XX 12-FEB-2004 (first entry)  
XX  
XX Bg8743c pneumococcal surface protein A (PspA) central region.  
DE  
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX immunological; gene therapy; immunostimulant.  
XX  
XX Unidentified.  
XX  
XX US6592876-B1.  
XX  
XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.  
XX  
XX 20-APR-1993; 93US-00048896.  
XX  
XX 06-JUN-1995; 95US-00465746.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
XX WPI; 2003-862841/80.  
XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 48; 121pp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Bg8743c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 166 AA;  
Query Match 82.0%; Score 393; DB 7; Length 166;  
Best Local Similarity 85.0%; Pred. No. 1.1e-28;  
Matches 85; Conservative 6; Mismatches 7; Indels 2; Gaps 2;  
QY 1 LKEIDESSEDYVKEGLRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEKVDFFK 60  
DB 1 LKEIDESSEDYVKEGLRAPLQSKLDKAKKLSKLELSKDIDELDAEIAKLEKVDGDFP 60  
QY 61 NSDGEQA-QYLAAAEEDL-AKKALEKTEADLKAVHEPE 98  
DB 61 NSDGEQAQYLVAAEKDLDAKEAEELGNTGADLKAVDEPE 100  
RESULT 12  
AAW14566  
ID AAW14566 standard; protein; 185 AA.  
XX  
XX AAW14566;  
XX  
XX 17-OCT-2003 (revised)  
XX  
XX 28-OCT-1997 (first entry)  
XX  
XX Streptococcus pneumoniae PspA central region.  
DE  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
XX bacteraemia; pneumonia.  
XX  
XX Streptococcus pneumoniae; strain Ac94.  
XX  
XX WO9709994-A1.  
XX  
XX 20-MAR-1997.  
XX  
XX 16-SEP-1996; 96WO-US014819.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 XX Example 6; Fig 13; 296pp; English.  
 PS  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 80.1%; Score 383.5; DB 2; Length 185;  
 Best Local Similarity 85.1%; Pred. No. 9.1e-28;  
 Matches 86; Conservative 4; Mismatches 8; Indels 3; Gaps 3;  
 QY 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 59  
 DB 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 60  
 QY 60 KNS-DGEQAOYLAAEEDL-AKKAELKTEADLKAVHEPE 98  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELKTEADLKAVNEPE 101  
 RESULT 13  
 ID AAW02600  
 AC AAW02600 standard; protein; 185 AA.  
 XX  
 XX AAW02600;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 XX US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 46; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 80.1%; Score 383.5; DB 7; Length 185;  
 Best Local Similarity 85.1%; Pred. No. 9.1e-28;  
 Matches 86; Conservative 4; Mismatches 8; Indels 3; Gaps 3;  
 QY 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 59  
 DB 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 60  
 QY 60 KNS-DGEQAOYLAAEEDL-AKKAELKTEADLKAVHEPE 98  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELKTEADLKAVNEPE 101  
 RESULT 14  
 ID AAW14571  
 AC AAW14571 standard; protein; 204 AA.  
 XX  
 XX AAW14571;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Ef1019.  
 XX  
 XX WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 XX Example 6; Fig 13; 296pp; English.  
 PS  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 80.1%; Score 383.5; DB 2; Length 185;  
 Best Local Similarity 85.1%; Pred. No. 9.1e-28;  
 Matches 86; Conservative 4; Mismatches 8; Indels 3; Gaps 3;  
 QY 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 59  
 DB 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 60  
 QY 60 KNS-DGEQAOYLAAEEDL-AKKAELKTEADLKAVHEPE 98  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELKTEADLKAVNEPE 101  
 RESULT 13  
 ID AAW02600  
 AC AAW02600 standard; protein; 185 AA.  
 XX  
 XX AAW02600;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 XX US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 46; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 80.1%; Score 383.5; DB 7; Length 185;  
 Best Local Similarity 85.1%; Pred. No. 9.1e-28;  
 Matches 86; Conservative 4; Mismatches 8; Indels 3; Gaps 3;  
 QY 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 59  
 DB 1 LKEIDESSEDYVKEGLRPLQSELDVAKQAKLKLSELSKIDELDAEIAK-LEKDVDF 60  
 QY 60 KNS-DGEQAOYLAAEEDL-AKKAELKTEADLKAVHEPE 98  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELKTEADLKAVNEPE 101

CC surface protein A (PspA) of Streptococcus pneumoniae strain Efi1019.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 204 AA;

Query Match 70.9%; Score 339.5; DB 2; Length 204;  
Best Local Similarity 73.7%; Pred. No. 1.3e-23;  
Matches 73; Conservative 10; Mismatches 15; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
DB 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
QY 61 NSDGEQAQYLAAAEEDL-AKKAELKTEADLKKAHVHEPE 98  
DB 61 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAHVHEPE 99

RESULT 15  
ABW02605  
ID ABW02605 standard; protein; 204 AA.  
XX  
AC ABW02605;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Efi1019c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
OS  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX

Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
PI WPI; 2003-862841/80.  
XX  
DR  
XX  
PT Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX

Example 6; SEQ ID NO 51; 121pp; English.  
XX  
XX  
CC The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Efi1019c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
XX used in the exemplification of the invention

SQ Sequence 204 AA;

Query Match 70.9%; Score 339.5; DB 7; Length 204;  
Best Local Similarity 73.7%; Pred. No. 1.3e-23;  
Matches 73; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
DB 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60  
QY 61 NSDGEQAQYLAAAEEDL-AKKAELKTEADLKKAHVHEPE 98  
DB 61 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAHVHEPE 99

Search completed: June 18, 2005, 16:51:18  
Job time : 72.6116 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.5405 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-1  
Perfect score: 479  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KKALEKTEADLKAVHEPE 98

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	98	US-09-147-875A-1	Sequence 1, Appli
2	461.5	96.3	99	US-08-710-749-9	Sequence 9, Appli
3	447	93.3	100	US-09-147-875A-5	Sequence 5, Appli
4	438	91.4	100	US-09-147-875A-2	Sequence 2, Appli
5	431	90.0	194	US-08-529-055-64	Sequence 64, Appli
6	431	90.0	550	US-09-583-110-4871	Sequence 4871, Ap
7	431	90.0	550	US-09-107-433-3858	Sequence 3858, Ap
8	431	90.0	8991	US-08-714-741-32	Sequence 32, Appli
9	430	89.8	100	US-09-147-875A-4	Sequence 4, Appli
10	429.5	89.7	101	US-08-710-749-4	Sequence 4, Appli
11	428	89.4	100	US-09-147-875A-3	Sequence 3, Appli
12	426	88.9	183	US-08-529-055-50	Sequence 50, Appli
13	424	88.5	168	US-08-529-055-55	Sequence 55, Appli
14	421	87.9	100	US-09-147-875A-6	Sequence 6, Appli
15	420.5	87.8	101	US-08-710-749-1	Sequence 1, Appli
16	419.5	87.6	101	US-08-710-749-2	Sequence 2, Appli
17	413.5	86.3	101	US-08-710-749-3	Sequence 3, Appli
18	410.5	85.7	101	US-08-710-749-5	Sequence 5, Appli
19	393	82.0	166	US-08-529-055-48	Sequence 48, Appli
20	392	81.8	100	US-09-147-875A-8	Sequence 8, Appli
21	390.5	81.5	101	US-09-147-875A-9	Sequence 9, Appli
22	383.5	80.1	185	US-08-529-055-46	Sequence 46, Appli
23	379.5	79.2	101	US-08-710-749-7	Sequence 7, Appli
24	378	78.9	100	US-09-147-875A-7	Sequence 7, Appli
25	373	77.9	102	US-08-710-749-8	Sequence 8, Appli
26	365.5	76.3	101	US-08-710-749-6	Sequence 6, Appli
27	339.5	70.9	99	US-08-710-749-10	Sequence 10, Appli

28	339.5	70.9	99	4	US-09-147-875A-11	Sequence 11, Appli
29	339.5	70.9	204	4	US-08-529-055-51	Sequence 51, Appli
30	331.5	69.2	170	4	US-08-529-055-60	Sequence 60, Appli
31	331.5	69.2	181	4	US-08-529-055-42	Sequence 42, Appli
32	331.5	69.2	864	4	US-08-714-741-40	Sequence 40, Appli
33	328.5	68.6	99	2	US-08-710-749-15	Sequence 15, Appli
34	328.5	68.6	188	4	US-08-529-055-59	Sequence 59, Appli
35	326.5	68.2	99	2	US-08-710-749-17	Sequence 17, Appli
36	324.5	67.7	99	4	US-09-147-875A-16	Sequence 16, Appli
37	324	67.6	100	4	US-09-147-875A-12	Sequence 12, Appli
38	323.5	67.5	99	2	US-08-710-749-11	Sequence 11, Appli
39	323.5	67.5	198	4	US-08-529-055-61	Sequence 61, Appli
40	323.5	67.5	619	1	US-08-465-746-2	Sequence 2, Appli
41	323.5	67.5	619	1	US-08-214-164-2	Sequence 3, Appli
42	323.5	67.5	619	2	US-08-467-852A-3	Sequence 2, Appli
43	323.5	67.5	619	2	US-08-246-636-2	Sequence 3, Appli
44	323.5	67.5	619	2	US-08-247-491A-3	Sequence 2, Appli
45	323.5	67.5	619	2	US-08-319-795-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-147-875A-1  
; Sequence 1, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-1

Query Match 100.0%; Score 479; DB 4; Length 98;  
Best Local Similarity 100.0%; Pred. No. 6.3e-37;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRAPLQSELDKQAKLSELSKIDELDAIKLEKQVEDFK 60  
DB 1 LKEIDSESDYVKEGLRAPLQSELDKQAKLSELSKIDELDAIKLEKQVEDFK 60  
QY 61 NSDGEQAYLAAREEDLAKKAELEKTEADLKAVHEPE 98  
DB 61 NSDGEQAYLAAREEDLAKKAELEKTEADLKAVHEPE 98

RESULT 2  
US-08-710-749-9  
; Sequence 9, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: 454312-2471
; APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-9

Query Match          96.3%; Score 461.5; DB 2; Length 99;
Best Local Similarity 98.0%; Pred. No. 2.5e-35;
Matches 97; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRAPLQSEL-DAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 59
Db 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 60

Qy 60 KNSDGEQAQYLAAAEEDLAKKAELEKTEADLKKAHVHEPE 98
Db 61 KNSDGEQAQYLAAAEEDLAKKAELEKTEADLKKAHVNEPE 99

RESULT 3
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match          93.3%; Score 447; DB 4; Length 100;
Best Local Similarity 96.0%; Pred. No. 5.4e-34;
Matches 96; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 60
Db 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 60

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAHVHEPE 98
Db 61 NSDGEQAQYLAAAEEDLAKKAELEKTEADLKKAHVHEPE 100

RESULT 4
US-09-147-875A-2
; Sequence 2, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
```

```
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-2

Query Match          91.4%; Score 438; DB 4; Length 100;
Best Local Similarity 94.0%; Pred. No. 3.6e-33;
Matches 94; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 60
Db 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKVEDF 60

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAHVHEPE 98
Db 61 NSNGEQAQYRAAAEEDLAAKQAELEKTEADLKKAHVHEPE 100

RESULT 5
US-08-529-055-64
; Sequence 64, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64
```

```

Query Match          90.0%; Score 431; DB 4; Length 194;
Best Local Similarity 93.9%; Pred. No. 3.4e-32;
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
Db 1 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 60

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEP 97
Db 61 XSDGEQAGQYLAARAEEDLAKKAELQTEADLKAVNEP 99

RESULT 6
US-09-583-110-4871
; Sequence 4871, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4871
; TYPE: PRT
; LENGTH: 550
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4871

Query Match          90.0%; Score 431; DB 4; Length 550;
Best Local Similarity 93.0%; Pred. No. 1.1e-31;
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
Db 144 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 203

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEP 98
Db 204 NSNGEQAQYRAAEEDLAKKAELKTEADLKAVNEP 243

RESULT 7
US-09-107-433-3858
; Sequence 3858, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <unknown>
; OPERATING SYSTEM: <unknown>
; SOFTWARE: <unknown>
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match          90.0%; Score 431; DB 4; Length 550;
Best Local Similarity 93.0%; Pred. No. 1.1e-31;
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
Db 144 LKEIDESDSEYVKEGLRAPLQSELDKAKLSKLEELSDKIDELDAETAKLEKVEDFK 203

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEP 98
Db 204 NSNGEQAQYRAAEEDLAKKAELKTEADLKAVNEP 243

RESULT 8
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF.
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741

```

```

; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match          90.0%; Score 431; DB 4; Length 8991;
Best Local Similarity 93.9%; Pred. No. 2.9e-30;
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy      1  LKEIDSDSDYVKEGLRAPLQSELDQAQAKSLKEELSDKIDELDAEIAKLEKVDFFK 60
      |||
      7537 LKEIDSDSDYVKEGLRAPLQSELDQAQAKSLKEELSDKIDELDAEIAKLEKVDFFK 7596

Db

Qy      61  NSDGEQA-QYLAAREDL-AKKALEKTEADLKKAVHEP 97
      |||
      7597 XSDGEQAQGYLAAREDLIAKKALEKTEADLKKAVNEP 7635

Db

```

```

RESULT 9
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

```

```

RESULT 10
US-08-710-749-4
; Sequence 4, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
JIS-08-710-749-4

```

RESULT 11
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

```

Query Match	89.4%	Score 428;	DB 4;	Length 100;
Best Local Similarity	92.0%;	Pred. No. 3e-32;		
Matches	92;	Conservative	4;	Mismatches 2; Indels 2; Gaps 2;
Qy	1	LKEIDESDSYVKEGLRAPLQSELDQAQKLSKLEELSDKIDELDAEIAKLEKQVDFK	60	
Db	1	LKEIDESDSYVKEGLRAPLQSELDQAQKLSKLEELSDKIDELDAEIAKLEKQVDFK	60	
Qy	61	NSDGEQA-QYLAABEDL-AKKAELEKTEADLKKAVHEP	98	



Db 61 NSNGEAEQYRAAGEDLAAQAELEKTEADLKAVHEPE 100

## RESULT 12

US-08-529-055-50

; Sequence 50, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Thereof, and Uses of Such Genes,

; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/529,055

; FILING DATE: 15-SEP-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2400

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 183 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-529-055-50

Query Match 88.9%; Score 426; DB 4; Length 183;

Best Local Similarity 92.0%; Pred. No. 9.1e-32;

Matches 92; Conservative 2; Mismatches 4; Indels 2; Gaps 2;

QY 1 LKEIDSESDYVKGRLAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

Db 1 LKEIDSESDYVKGRLAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98

Db 61 NSDGEQAQYLAAGEDLAAQAELEKTEADLKAVDEPE 100

## RESULT 13

US-08-529-055-55

; Sequence 55, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Thereof, and Uses of Such Genes,

; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/529,055

; FILING DATE: 15-SEP-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2400

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 55:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 168 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-529-055-55

Query Match 88.5%; Score 424; DB 4; Length 168;

Best Local Similarity 92.0%; Pred. No. 1.3e-31;

Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1 LKEIDSESDYVKGRLAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

Db 1 LKEIDSESDYVKGRLAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98

Db 61 NSDGEQAQYLAAGEDLAAQAELEKTEADLKAVDEPE 100

## RESULT 14

US-09-147-875A-6

; Sequence 6, Application US/09147875A

; Patent No. 6638516

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/09/147,875A

; CURRENT FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 100

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (1)..(100)

; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid

US-09-147-875A-6

Query Match 87.9%; Score 421; DB 4; Length 100;

Best Local Similarity 92.0%; Pred. No. 1.3e-31;  
Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKQVDFK 60  
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKQVDFK 60

Qy 61 NSDGEQA-QYLAAAEEDL-AKKALEKTEADLKAVHEPE 98  
Db 61 NSDGEQAQYLAIAAEEDLAKKAXLEKAEADLKAVDEPE 100

RESULT 15  
US-08-710-749-1  
; Sequence 1, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-1

Query Match 87.8%; Score 420.5; DB 2; Length 101;  
Best Local Similarity 92.1%; Pred. No. 1.5e-31;  
Matches 93; Conservative 4; Mismatches 1; Indels 3; Gaps 3;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKQVDFK 59  
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKQVDFK 60

Qy 60 KNSDGEQA-QYLAAAEEDL-AKKALEKTEADLKAVHEPE 98  
Db 61 KNSNGEQAQYRAAAAEEDLAAKQAELEKTEADLKAVNEPE 101

Search completed: June 18, 2005, 17:07:04  
Job time : 19.5405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 61.7037 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-1  
Perfect score: 479  
Sequence: 1 LKEIDSESDYVVEKGLRAP.....KKAELEKTEADLKKAHVEPE 98

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10F\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	98	US-10-674-755-1	Sequence 1, Appli
2	447	93.3	100	US-10-674-755-1	Sequence 5, Appli
3	438	91.4	100	US-10-674-755-2	Sequence 2, Appli
4	431	90.0	194	US-10-299-636-79	Sequence 79, Appli
5	430	89.8	100	US-10-674-755-4	Sequence 3, Appli
6	428	89.4	100	US-10-674-755-3	Sequence 4, Appli
7	426	88.9	183	US-10-299-636-65	Sequence 65, Appli
8	424	88.5	168	US-10-299-636-70	Sequence 70, Appli
9	421	87.9	100	US-10-674-755-6	Sequence 6, Appli
10	393	82.0	166	US-10-299-636-63	Sequence 63, Appli
11	392	81.8	100	US-10-674-755-8	Sequence 8, Appli

12	390.5	81.5	101	15	US-10-674-755-9	Sequence 9, Appli
13	383.5	80.1	185	15	US-10-299-636-61	Sequence 61, Appli
14	378	78.9	100	15	US-10-674-755-7	Sequence 7, Appli
15	339.5	70.9	99	15	US-10-674-755-11	Sequence 11, Appli
16	339.5	70.9	204	15	US-10-299-636-66	Sequence 66, Appli
17	331.5	69.2	170	15	US-10-299-636-75	Sequence 75, Appli
18	331.5	69.2	181	15	US-10-299-636-57	Sequence 57, Appli
19	331.5	69.2	643	15	US-10-299-636-95	Sequence 95, Appli
20	331.5	69.2	670	9	US-09-748-875-63	Sequence 63, Appli
21	331.5	69.2	670	10	US-09-298-523B-63	Sequence 63, Appli
22	331.5	69.2	690	9	US-09-748-875-61	Sequence 61, Appli
23	331.5	69.2	690	10	US-09-298-523B-61	Sequence 61, Appli
24	331.5	69.2	691	9	US-09-748-875-1	Sequence 1, Appli
25	331.5	69.2	691	10	US-09-298-523B-1	Sequence 1, Appli
26	331.5	69.2	701	9	US-09-748-875-62	Sequence 62, Appli
27	331.5	69.2	701	10	US-09-298-523B-62	Sequence 62, Appli
28	331.5	69.2	707	9	US-09-748-875-2	Sequence 2, Appli
29	331.5	69.2	707	10	US-09-298-523B-2	Sequence 2, Appli
30	331.5	69.2	711	9	US-09-748-875-3	Sequence 3, Appli
31	331.5	69.2	711	10	US-09-298-523B-3	Sequence 3, Appli
32	331.5	69.2	739	17	US-10-732-923-3294	Sequence 3294, Ap
33	331.5	69.2	929	9	US-09-748-875-60	Sequence 60, Appli
34	331.5	69.2	929	10	US-09-298-523B-60	Sequence 60, Appli
35	331.5	69.2	929	15	US-10-299-636-94	Sequence 94, Appli
36	328.5	68.6	188	15	US-10-299-636-74	Sequence 74, Appli
37	324.5	67.7	99	15	US-10-674-755-16	Sequence 16, Appli
38	324	67.6	100	15	US-10-674-755-12	Sequence 12, Appli
39	323.5	67.5	198	15	US-10-299-636-76	Sequence 76, Appli
40	323.5	67.5	354	15	US-10-299-636-105	Sequence 105, App
41	323.5	67.5	588	15	US-10-299-636-36	Sequence 36, Appli
42	323.5	67.5	619	10	US-09-882-774-1	Sequence 1, Appli
43	323.5	67.5	619	15	US-10-282-122A-73702	Sequence 73702, A
44	323.5	67.5	619	16	US-10-414-532-72	Sequence 72, Appli
45	320	66.8	100	15	US-10-674-755-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1  
US-10-674-755-1  
; Sequence 1, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-1

Query Match 100.0%; Score 479; DB 15; Length 98;  
Best Local Similarity 100.0%; Pred. No. 3.2e-30;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	LKEIDSESDYVVEKGLRAPLQSELDKAKLSELSDKIDELDAITAKLEKDVDPK 60
Db	1	LKEIDSESDYVVEKGLRAPLQSELDKAKLSELSDKIDELDAITAKLEKDVDPK 60
Qy	61	NSDGEAQYLAABEDLAKKAELEKTEADLKKAHVEPE 98
Db	61	NSDGEAQYLAABEDLAKKAELEKTEADLKKAHVEPE 98

RESULT 2

```
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match          93.3%; Score 447; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 1e-27;
Matches 96; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQA-QYLAARAEEDL-AKKAELKTEADLKKAVHEPE 98
   |||||
Db 61 NSDGEQAQYLAARAEEDLIAKKAELQTEADLKKAVHEPE 100

RESULT 3
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match          91.4%; Score 438; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 5.2e-27;
Matches 94; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQA-QYLAARAEEDL-AKKAELKTEADLKKAVHEPE 98
   |||||
Db 61 NSNGEQAQYRAARAEEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 4
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
```

```
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match          90.0%; Score 431; DB 15; Length 194;
Best Local Similarity 93.9%; Pred. No. 3.7e-26;
Matches 93; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 1 LKEIDSDSDYVKEGFAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQA-QYLAARAEEDL-AKKAELKTEADLKKAVHEP 97
   |||||
Db 61 XSDGEQAQYLAARAEEDLIAKKAELQTEADLKKAVNEP 99

RESULT 5
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match          89.8%; Score 430; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 2.2e-26;
Matches 94; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQA-QYLAARAEEDL-AKKAELKTEADLKKAVHEP 98
   |||||
Db 61 NSDGEQAQYLAARAEEDLIAKKAELKTEADLKKAVDEP 100

RESULT 6
US-10-674-755-3
```

; Sequence 3, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-3

Query Match 89.4%; Score 428; DB 15; Length 100;  
Best Local Similarity 92.0%; Pred. No. 3.1e-26;  
Matches 92; Conservative 4; Mismatches 2; Indels 2; Gaps 2;  
QY 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFK 60  
DB 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFK 60  
QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEPE 98  
DB 61 NSNGEAEQYRAAAGDAAKQAELEKTEADLKKAVHEPE 100

RESULT 7  
US-10-299-636-65  
; Sequence 65, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 65  
; LENGTH: 183  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-65

Query Match 88.9%; Score 426; DB 15; Length 183;  
Best Local Similarity 92.0%; Pred. No. 8.6e-26;  
Matches 92; Conservative 2; Mismatches 4; Indels 2; Gaps 2;  
QY 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFK 60  
DB 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFQ 60  
QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEPE 98  
DB 61 NSDGEQAQYLAAGDAAKQAELEKTEADLKKAVHEPE 100

RESULT 8  
US-10-299-636-70  
; Sequence 70, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 168  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
; NAME/KEY: UNSURE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa at position 38 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (42)  
; OTHER INFORMATION: Xaa at position 42 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (84)  
; OTHER INFORMATION: Xaa at position 84 is unknown  
US-10-299-636-70

Query Match 88.5%; Score 424; DB 15; Length 168;  
Best Local Similarity 92.0%; Pred. No. 1.1e-25;  
Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;  
QY 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFK 60  
DB 1 LKEIDESDSEYVKEGLRPLQSELDAAKQAKSLKEELSDKIDELDAETAKLEKQVDFK 60  
QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAVHEPE 98  
DB 61 NSDGEQAQYLAAGDAAKQAELEKTEADLKKAVHEPE 100

RESULT 9  
US-10-674-755-6  
; Sequence 6, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 100  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:

```
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match      87.9%; Score 421; DB 15; Length 100;
Best Local Similarity 92.0%; Pred. No. 1.1e-25;
Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDKAKQAKLSKLELSKIDELDAEIAKLEKVDFFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYVKEGERAPLQSELDKAKQAKLSKLEKSKDKXDELDAEIAKLEKVDFFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAHVEPE 98
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAQYLAIAAEEDLIAKKALEKAEADLKKAHVEPE 100

RESULT 10
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      82.0%; Score 393; DB 15; Length 166;
Best Local Similarity 85.0%; Pred. No. 2.9e-23;
Matches 85; Conservative 6; Mismatches 7; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDKAKQAKLSKLELSKIDELDAEIAKLEKVDFFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSDSDYIKEGRLAPLQSKLDKAKKAKLSKIDELSKIDELDAEIAKLEKVDGDFP 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAHVEPE 98
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAQYLAIAAEEDLIAKKALEKAEILGNTGADLKKAHVEPE 100

RESULT 11
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-8

Query Match      81.8%; Score 392; DB 15; Length 100;
Best Local Similarity 86.0%; Pred. No. 2e-23;
Matches 86; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDKAKQAKLSKLELSKIDELDAEIAKLEKVDFFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKGIDSDSDYVKEGLRAPLQSELDKAKRKYKLTSTLELSKIDELDAEIPKLEKNVYFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKKAHVEPE 98
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LTDAEQTEQYLAIAAEKDLADKKALEKTEADLKKAHVEPE 100

RESULT 12
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      81.5%; Score 390.5; DB 15; Length 101;
Best Local Similarity 86.1%; Pred. No. 2.6e-23;
Matches 87; Conservative 3; Mismatches 8; Indels 3; Gaps 3;

Qy 1 LKEIDSDSDYVKEGLRAPLQSELDKAKQAKLSKLELSKIDELDAEIAK-LEKVDFF 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSDSDYVKEGLRVPLQSELDVYKQAKLLELSKIDELDAEIAKLNKKVDFF 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 60 KNS-DGQQAQYLAIAAEEDL-AKKAELEKTEADLKKAHVEPE 98
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 QNSGGYSALYLEAAEKDLVAKKALEKTEADLKKAHVEPE 101
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
```

```

; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 61
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-61

```

```

Query Match      80.1%; Score 383.5; DB 15; Length 185;
Best Local Similarity 85.1%; Pred. No. 1.8e-22;
Matches 86; Conservative 4; Mismatches 8; Indels 3; Gaps 3;

QY      1 LXEIDSDSDYKGEGLRAPQLQSELDQAQKLSKLELSDKIDELDAEIAK-LEKQVDF 59
          |||||
Db       1 LXEIDSDSDYKGEGLRVPLQSELDVQAKLKLLELSDKIDELDAEIAKLNKQVDF 60
          |||||

QY      60 KNS-DGEQGYLAAREDL-AKKAELEKTEADLKVAHEPE 98
          |||||
Db       61 QNSGGYGYALYLEAAEKDLWAKKAELEKTEADLKVAHNEPE 101
          |||||

```

```

RESULT 14
US-10-674-755-7
; Sequence 7, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-7

```

	Query Match	78.9%;	Score 378;	DB 15;	Length 100;
	Best Local Similarity	83.0%;	Pred. No. 2.5e-22;		
	Matches	83;	Conservative 5;	Mismatches 10;	Indels 2; Gaps 2;
Qy	1	LKEIDESSEDYVVEGLRAPLQSELDKQAOKLSKLELSOKIDELDAEIAKLEKVDFFK	60		
Db	1	LKEIDESSEDYEVEGLRAPLQSKLDAAKAKLSKLEDSKDKDELDAEIAKLEKVDGDFP	60		
Qy	61	NSDGEQA-QYLAAAEEDL-AKKAELEKTEADLKAVHEPE	98		
Db	61	NSDGEQAQYLVAAEKOLDAAKEALGNTGADLKAVDPE	100		

```

RESULT 15
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1993-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1

```

```
; SEQ ID NO 11
; LENGTH: 99
; TYPE: prt
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11
```

```
Query Match          70.9%; Score 339.5; DB 15; Length 99;
Best Local Similarity 73.7%; Pred. No. 2.5e-19;
Matches 73; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

QY      1 LKEIDSDSDYKGEGRAPQLQSGLDAKQAOKLSKLEELSDKIDELDAEIAKLEKDVEDFK 60
        |||||||
Db       1 LKEIDSDSDYKGEGRAPQLQSGLDAKQAOKLSKLEELSDKIDELDAEIAKLEKDVEDFK 60
        |||||||

QY      61 NSDGEQAQYLAAAEEDL-AKKAELEKTEADLKAVHEPE 98
        :::::|-----:|||
Db       61 ENNNVDYFKEGLEKTIAAKKAELEKTEADLKAVNEPE 99
        :::::|-----:|||
```

Search completed: June 18, 2005, 18:00:20  
Job time : 62.7037 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 12.7528 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-1  
Perfect score: 479  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KKAELEKTEADLKKAVHEPE 98

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	323.5	67.5	619	2 A97887	surface protein ps
2	323.5	67.5	619	2 A41971	surface protein ps
3	121	25.3	744	2 F95013	pneumococcal surfa
4	107	22.3	896	2 S43074	epidermal growth f
5	107	22.3	1319	2 A28313	glued protein - fr
6	106	22.1	161	2 S48396	tropomyosin TPM2 -
7	106	22.1	1269	2 F84730	probable myosin he
8	104	21.7	886	2 H69378	conserved hypothet
9	104	21.7	3488	2 T34418	hypothetical prote
10	103.5	21.6	784	2 T05409	hypothetical prote
11	102	21.3	1006	2 C70445	ATPase subunit of
12	102	21.3	1110	2 I51116	NF-180 - sea lamp
13	100.5	21.0	1169	2 A64505	p115 homolog - Met
14	100	20.9	281	2 F75216	hypothetical prote
15	100	20.9	924	2 S06117	myosin heavy chain
16	99.5	20.8	408	2 S30283	protein M precursor
17	99.5	20.8	603	2 T00379	KIAA0640 protein -
18	99.5	20.8	1027	2 S37711	kinesin heavy chain
19	99	20.7	347	2 JCS788	tsec-1 protein (A
20	99	20.7	2007	1 B43402	myosin heavy chain
21	98.5	20.6	554	2 A60115	M protein precursor
22	98.5	20.6	897	2 A54696	EGF receptor subst
23	98	20.5	1976	2 A59252	myosin heavy chain
24	98	20.5	2288	2 T29999	hypothetical prote
25	97.5	20.4	388	2 A49545	plasmaogen-bindin
26	97	20.3	1138	2 T24635	hypothetical prote
27	97	20.3	1156	2 B70356	chromosome assembl
28	96.5	20.1	436	2 S30284	M protein precursor
29	96.5	20.1	472	2 S43554	plasmaogen-bindin

30 96.5 20.1 1790 2 S67593 transport protein  
31 96 20.0 522 2 G02533 occludin - human  
32 95.5 19.9 532 2 S54871 M protein - Strept  
33 95.5 19.9 562 2 G70002 hypothetical prote  
34 95.5 19.9 1053 2 A41642 dynactin - chicken  
35 95.5 19.9 1300 2 I53799 CGI protein - huma  
36 95.5 19.9 1356 2 S32763 kinectin 1 - human  
37 95.5 19.9 1927 2 A59236 embryonic muscle m  
38 94.5 19.7 483 2 A26297 M6 protein - Strept  
39 94.5 19.7 484 2 S46489 M1 protein precurs  
40 94.5 19.7 501 2 A44643 M protein precurs  
41 94.5 19.7 880 2 F75103 conserved hypothet  
42 94.5 19.7 1938 1 A40997 myosin heavy chain  
43 94.5 19.7 1940 2 A29320 myosin heavy chain  
44 94 19.6 1392 2 A43336 microtubule-vesicl  
45 94 19.6 1427 2 S22695 restin - human

ALIGNMENTS

RESULT 1

A97887 surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
y, P.; Sun, E.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:G  
C:Genetics:  
A:Gene: pspA

Query Match 67.5%; Score 323.5; DB 2; Length 619;  
Best Local Similarity 69.7%; Pred. No. 1.5e-14;  
Matches 69; Conservative 13; Mismatches 16; Indels 1; Gaps 1;  
QY 1 LKEIDSESDYVKEGLRAPLQSELDKAKLSEELSDKIDELDAETAKLEKVEDPK 60  
DB 223 LKEIDSESDYVKEGLRAPLQSELDKAKLSEELSDKIDELDAETAKLEKVEDPK 282  
QY 61 NSDGEQAOYLAAEEDL-AKKALEKTEADLKKAVHEPE 98  
DB 283 ENNVEDYFKEGLEKTIKKALEKTEADLKKAVNEPE 321

RESULT 2

A41971 surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA270  
A:Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Yoellinger, D.C.; Yother, J.; Briles, D.E.

```
Infect. Immun. 59, 1285-1289, 1991
A:Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A:Reference number: A60282; MUID:91169598; PMID:2004810
A:Accession: A60282
A:Molecule type: protein
A:Residues: 32-76 <TAL>
A:Experimental source: strain JY2008
C:Genetics:
A:Gene: pspA
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-619/Product: surface protein pspA #status predicted <MAT>
F:411-430/Domain: cpl repeat homology <CP01>
F:431-450/Domain: cpl repeat homology <CP02>
F:451-470/Domain: cpl repeat homology <CP03>
F:471-490/Domain: cpl repeat homology <CP04>
F:491-510/Domain: cpl repeat homology <CP05>
F:511-530/Domain: cpl repeat homology <CP06>
F:531-550/Domain: cpl repeat homology <CP07>
F:551-570/Domain: cpl repeat homology <CP08>
F:571-591/Domain: cpl repeat homology <CP09>
F:592-611/Domain: cpl repeat homology <CP10>
Query Match 67.5%; Score 323.5; DB 2; Length 619;
Best Local Similarity 69.7%; Pred. No. 1.5e-14;
Matches 69; Conservative 13; Mismatches 16; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYVKEGLRAPLQSELDKQAKLSKLELSDKIDELDAIAKLEKDVDFK 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 LKEIDSESDYAKGFGFAPLQSKLDKAKKLSKLELSDKIDELDAIAKLEKDLQKAAE 282
Qy 61 NSDGEFAQYLAAREDL-AKKAELEKTEADLKAVHEPE 98
Db :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
283 ENNVEDYFKGLEKTIIAAKKAELEKTEADLKAVNEPE 321
RESULT 3
F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95013
R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95013
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-744 <KUR>
A:Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:G
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0117
Query Match 25.3%; Score 121; DB 2; Length 744;
Best Local Similarity 33.3%; Pred. No. 0.52;
Matches 44; Conservative 22; Mismatches 24; Indels 42; Gaps 7;
Qy 2 KEIDE-----SDSEDYVKEGLRAPLQSELDKQAKLSK-----LEELSDKI-----D 43
Db 314 KEISNLEILLGGADPEDT-----AALQNKLAAKKAELAKKQTELEKLLSLDPGKTD 368
Qy 44 ELD--AEIAKLEKVEDFKNSDGEFAQYL-----AAAEEDLA-----KKAELE 84
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 ELDKAEAEELDKADELQNKVADLEKISNLEILLGGADSEDDTAALQNKLYAKKAELE 428
Qy 85 KTEADLKAVHE 96
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
429 KTKELDAALNE 440
```

RESULT 4

```
S43074
epidermal growth factor receptor substrate - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43074; I38525
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.
Oncogene 9, 1039-1045, 1994
A:Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4,
A:Reference number: S43074; MUID:94181254; PMID:8134107
A:Accession: S43074
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-896 <BER>
A:Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:9470034; PIDN:CAA82305.1; PID:94700
R:Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner,
Oncogene 9, 1591-1597, 1994
A:Title: The human epi15 gene, encoding a tyrosine kinase substrate, is conserved in evo
A:Reference number: I38525; MUID:94239734; PMID:8193552
A:Accession: I38525
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-821, 'M', 823-896 <REG>
A:Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260
C:Genetics:
A:Gene: GDB:EPF15; AF-1P; MLLT5
A:Cross-references: GDB:360337; OMIM:600051
A:Map position: 1p32-1p32
```

Query Match 22.3%; Score 107; DB 2; Length 896;

Best Local Similarity 26.7%; Pred. No. 5.4; Matches 31; Conservative 28; Mismatches 33; Indels 24; Gaps 4;

Qy 3 EIDSESDYVKEGLR--APLQSELDKQAKLSKL-----EELSDKIDELDAIAKLEKDV 56

Db 353 EQDLKEKEDITKQRTSEVQDLQDEVQRENTNLQKQAKQVQVQLDELDEQKAEQQL 412

Qy 57 EDFKNSDGEFAQYLAAR-----EEDLAKKAE-----LEKTEADLKAV 94

Db 413 KEVRKKAEBEAQITSSILKABLTQSQISTYEELAKARELSLQOETALEESV 468

RESULT 5

```
A28313
glued protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A28313
R:Swaroop, A.; Swaroop, M.; Garen, A.
Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987
A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ge
A:Reference number: A28313; MUID:87317680; PMID:2819881
A:Accession: A28313
A:Molecule type: DNA; mRNA
A:Residues: 1-1319 <SWA>
A:Cross-references: UNIPROT:P13496
A:Note: the authors' translation is inconsistent with the nucleotide sequence in the reg
C:Genetics:
A:Gene: FlyBase:Gl
A:Cross-references: FlyBase:FBgn0001108
A:Introns: 18/2; 479/3
C:Keywords: cytoskeleton; glycoprotein
F:397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #8
```

Query Match 22.3%; Score 107; DB 2; Length 1319;

Best Local Similarity 29.2%; Pred. No. 7.9; Matches 33; Conservative 25; Mismatches 33; Indels 22; Gaps 4;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDKQAKLSKL-----EELSDKIDELDAIAKLEKDV 56

Db 429 LRLDSAHDKHDIQK-----LSKLEMKRSEVTELTKEKLSAKIDELEIVADLQEQV 482

QY 57 EDFNSDGEQAQVLA-----AABEDLAKAELEKTEADLKKAVHEPE 98  
DB 483 DAALGAE-EMVEQLAEKMELEDKVLLBEETAQLEALBEVHQLVESNHELE 534

RESULT 6  
S48396  
tropomyosin TPM2 - yeast (Saccharomycetes cerevisiae)  
N:Alternate names: protein YIL138c  
C:Species: Saccharomycetes cerevisiae  
C>Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 09-Jul-2004  
C:Accession: S48396; A56490  
R:Churcher, C.  
submitted to the EMBL Data Library, September 1994  
A:Reference number: S48310  
A:Accession: S48396  
A:Molecule type: DNA  
A:Residues: 1-161 <CHU>  
A:Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:G603997; PID:G763208; MI  
R:Drees, B.; Brown, C.; Barrell, B.G.; Bretscher, A.  
J. Cell Biol. 128, 383-392, 1995  
A:Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform disti  
A:Reference number: A56490; MUID:95146545; PMID:7844152  
A:Accession: A56490  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-161 <DRE>  
A:Cross-references: GB:Z47047; GB:Z38059; NID:G603997; PID:G763208  
C:Genetics:  
A:Gene: SGD:TPM2  
A:Cross-references: SGD:S0001400; MIPS:YIL138c  
A:Map position: 9L  
C:Superfamily: tropomyosin TPM1  
C:Keywords: cytoskeleton

Query Match 22.1%; Score 106; DB 2; Length 161;  
Best Local Similarity 34.3%; Pred. No. 1.1;  
Matches 34; Conservative 18; Mismatches 39; Indels 8; Gaps 3;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLESDKIDELDAETAKLEKVEDPK 60  
DB 11 LKLESESWOEK--HELRQKELEQSNTERKENEIKLSARNEQLDSEVEKLESQSDTK 68

QY 61 NSDGEQAQVLAABEDLAKK-----AELEKTEADLKKAV 94  
DB 69 QL-AEDSNLNENNTYTKNQDLEQLEDSEAKLKEAM 106

RESULT 7  
F84730  
Probable myosin heavy chain [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84730  
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84730  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1269 <STO>  
A:Cross-references: GB:AE002093; NID:G6598483; PIDN:AAC69932.2; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g32240  
A:Map position: 2

Query Match 22.1%; Score 106; DB 2; Length 1269;  
Best Local Similarity 33.0%; Pred. No. 8.8;  
Matches 33; Conservative 17; Mismatches 36; Indels 14; Gaps 3;

QY 6 ESDSEYVKEGLRAPLQSELDKQAKLSKLE-----ELSKIDELDAETAKLEK 54  
DB 661 EADSKYL--GQVAELQSTLEAFQVKKSSLEALNATENEKELTENLNAVTSKKKLEA 718

QY 55 DYEDFNSDGEQAQVLAABEDL-AKKALEKTEADLKKAA 93  
DB 719 TVDEYSVKISESENLESIRNELNVTQGLKESIENDLKA 758

RESULT 8  
H69378  
conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: H69378  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A:Reference number: A6250; MUID:98049343; PMID:9389475  
A:Accession: H69378  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-886 <KLE>  
A:Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:G2689355; PIDN:AB9021  
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 21.7%; Score 104; DB 2; Length 886;  
Best Local Similarity 27.0%; Pred. No. 8.4;  
Matches 37; Conservative 22; Mismatches 34; Indels 44; Gaps 4;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLESDKIDELDAETAKLEKVEDPK 60  
DB 296 LSEINQALRDVEKREG---DLTREAAGIQAKLKAEDNSKLEETIKRIEELERLEPFS 352

QY 61 NS-----DGEQAQV--LAAAEED-----LAK 79  
DB 353 KSHRLLETLPKMDRMQGIKAKLEKNLTPDKVKRYDLSKAKEEKEITEKLLKLIAR 412

QY 80 KAELEKTEADLKKAVHE 96  
DB 413 KSSLTRGALKKAAVEE 429

RESULT 9  
T34418  
hypothetical protein F12F3.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34418  
R:Fulton, B.; Wohldmann, P.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of C. elegans cosmid F12F3.  
A:Reference number: Z21521  
A:Accession: T34418  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3488 <FUL>  
A:Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN000023; CESP:F12F3.3  
A:Experimental source: strain Bristol N2; clone F12F3  
C:Genetics:  
A:Gene: CESP:F12F3.3  
A:Map position: 5  
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 21.7%; Score 104; DB 2; Length 3488;  
Best Local Similarity 37.7%; Pred. No. 33;  
Matches 43; Conservative 15; Mismatches 32; Indels 24; Gaps 7;

QY 2 KEIDES---DSEYVKEGLRAPLQSELDKQAKLSKL-----EELSKIDELDAEI 49



submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A;Reference number: A75001  
A;Accession: F75216  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-281 <KAW>  
A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CAB4918  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB2181

Query Match 20.9%; Score 100; DB 2; Length 281;  
Best Local Similarity 27.7%; Pred. No. 4.9;  
Matches 31; Conservative 30; Mismatches 31; Indels 20; Gaps 4;  
QY 1 LKEIDSDSE-DVYKGLRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEKQVEDF 59  
DB 133 LKYEYKLGQFEFEVAKIEA---AELESLEKAKKEIEELKGVKEKLEQEKLEKKL--- 186  
QY 60 KNSDGEQAOYLAAA-----EEDLAKKAELEKTEADLKKAHVHEPE 98  
DB 187 KESEVKLMVEYAKAKRAAELEAKLREYEEKVKEEELERKVSLESLNEYE 238

## RESULT 15

S06117  
myosin heavy chain, nonmuscle (clone lambda-FMHC) - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: S06117  
R;Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.  
Eur. J. Biochem. 184, 611-616, 1989  
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in  
s.  
A;Reference number: S06116; MUID:90032648; PMID:2806244  
A;Accession: S06117  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-924 <KAT>  
A;Cross-references: UNIPROT:Q02015; GB:X17590  
A;Note: this translation is not annotated in GenBank entry GGMHCFC, release 114  
C;Superfamily: myosin heavy chain; myosin motor domain homology  
F;1-303/Domain: myosin motor domain homology (fragment) <MMOT>

Query Match 20.9%; Score 100; DB 2; Length 924;  
Best Local Similarity 26.1%; Pred. No. 16;  
Matches 29; Conservative 24; Mismatches 34; Indels 24; Gaps 3;  
QY 1 LKEIDSDSE-DVYKGLRAPLQSELDKQAKLSKLEE-----LSDKI 42  
DB 484 IQDLLEQLDSE---EGARQKQLKQVSTEAKIKQMEIEILLLEDQNSKFLKEKLMEDRI 540  
QY 43 DELDAEIAKLEKQVEDFKNSDGEQAOYLAAAEEDLAKKAELEKTEADLKKA 93  
DB 541 AECTSQLAEIEEKAKNKLAKNKNQEMMTDLERLKE---EKTRQELEKA 588

Search completed: June 18, 2005, 17:03:49  
Job time : 13.8528 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 59.7417 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-1

Perfect score: 479

Sequence: 1 LKEIDSESDYVKEGLRAP.....KKAELEKTEADLKKAHPE 98

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_treml.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	443	92.5	406	2	Q9LAZ0 streptococc
2	441	92.1	340	2	Q8KQK5
3	440	91.9	394	2	Q9LAY6 streptococc
4	440	91.9	395	2	Q9LAZ1 streptococc
5	432	90.2	225	2	Q9L591 streptococc
6	431	90.0	222	2	Q9L577 streptococc
7	431	90.0	262	2	Q9L576 streptococc
8	431	90.0	415	2	Q9LAY7 streptococc
9	424	88.5	416	2	Q9LAY8 streptococc
10	423	88.3	194	2	Q9L5B5 streptococc
11	423	88.3	218	2	Q6UEB2 streptococc
12	423	88.3	233	2	Q9L568 streptococc
13	423	88.3	236	2	Q9L569 streptococc
14	423	88.3	243	2	Q9L564 streptococc
15	423	88.3	243	2	Q9L567 streptococc
16	423	88.3	244	2	Q9L565 streptococc
17	423	88.3	246	2	Q9L578 streptococc
18	423	88.3	247	2	Q9L566 streptococc
19	423	88.3	249	2	Q9L570 streptococc
20	423	88.3	254	2	Q9L563 streptococc
21	423	88.3	401	2	Q9LAZ2 streptococc
22	422	88.1	255	2	Q9L581 streptococc
23	422	88.1	255	2	Q9L5B6 streptococc
24	402	83.9	393	2	Q9LAZ3 streptococc
25	393	82.0	237	2	Q9L592 streptococc
26	393	82.0	395	2	Q9LAY9 streptococc
27	389	81.2	207	2	Q8GNS9 streptococc
28	339.5	70.9	417	2	Q9LAY3 streptococc
29	331.5	69.2	739	2	Q9RQT4 streptococc
30	331.5	69.2	820	2	Q9RQT1 streptococc
31	331.5	69.2	929	2	Q9KK19 streptococc

32 331.5 69.2 929 2 Q9ZAY5 streptococc  
33 323.5 67.5 619 2 Q54972 streptococc  
34 323.5 67.5 619 2 Q8DRI0 streptococc  
35 319.5 66.7 99 2 Q8KQK4 streptococc  
36 317.5 66.3 415 2 Q9LAY1 streptococc  
37 315.5 65.9 437 2 Q9LAY4 streptococc  
38 312.5 65.2 249 2 Q9L575 streptococc  
39 309.5 64.6 426 2 Q9L575 streptococc  
40 305.5 63.8 224 2 Q8GNS8 streptococc  
41 300.5 62.7 395 2 Q9LAY2 streptococc  
42 300.5 62.7 408 2 Q9LAY0 streptococc  
43 296 61.8 869 2 Q9KK27 streptococc  
44 164 34.2 479 2 Q9LAX2 streptococc  
45 164 34.2 480 2 Q9LAX3 streptococc

## ALIGNMENTS

### RESULT 1

Q9LAZ0 PRELIMINARY; PRT; 406 AA.  
AC Q9LAZ0;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]\_SEQUENCE FROM N.A.  
RC STRAIN=DL6A;  
RX MEDLINE=20448953; PubMed=10992499;  
RX DOI=10.1128/JAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of S.p.K. mosaic genes and evidence for past recombination in Streptococcus pneumoniae";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071805; AAF27701.1; -  
DR InterPro; IPR009082; His\_kin\_homodim.  
DR PRINTS; IPR00533; Tropomyosin.  
FT NON\_TER 406 406  
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D0F052 CRC64;

Query Match 92.5%; Score 443; DB 2; Length 406;  
Best Local Similarity 95.0%; Pred. No. 5,1e-20;  
Matches 95; Conservative 2; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDSESDYVKEGLRAPLQSELDKAKLSELSKLELSDKIDELDAIAKLEKQVEDPK 60  
DB 213 LKEIDSESDYVKEGLRAPLQSELDKAKLSELSKLELSDKIDELDAIAKLEKQVEDPK 272

QY 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKKAHPE 98  
DB 273 NSDGEQAQYLAAAEEDLAKKAELKTEADLKKAHPE 312

### RESULT 2

Q8KQK5 PRELIMINARY; PRT; 340 AA.  
AC Q8KQK5;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Pneumococcal surface protein A (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
RD DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082387; AAL92492.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1
FT NON TER 340
FT NON TER 340
SQ SEQUENCE 340 AA; 38023 MW; E507BCF0081FBD57 CRC64;

Query Match 92.1%; Score 441; DB 2; Length 340;
Best Local Similarity 95.0%; Pred. No. 5.8e-20;
Matches 95; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60
Db 197 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 256

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98
Db 257 NSDGEQAQYLAAAEEDLVAKKAELKTEADLKAVNEPE 296

RESULT 3
Q9LAY6 PRELIMINARY; PRT; 394 AA.
AC Q9LAY6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L81905;
RX MEDLINE=20448953; PubMed=10992499;
RD DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27705.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 394
FT NON TER 394
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C98FAA CRC64;

Query Match 91.9%; Score 440; DB 2; Length 394;
Best Local Similarity 95.0%; Pred. No. 7.6e-20;
Matches 95; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60
Db 213 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 272

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98
Db 273 NSDGEQAQYLAAAEEDLVAKKAELKTEADLKAVNEPE 312

RESULT 5
Q9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RD Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RD Beall B.W.;
RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
```

```

Db 273 NSDGEQAQYLAAAEEDLIAKKAELKAEADLKAVDEPE 312

RESULT 4
Q9LAZ1 PRELIMINARY; PRT; 395 AA.
AC Q9LAZ1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9739;
RX MEDLINE=20448953; PubMed=10992499;
RD DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27705.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395
FT NON TER 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECA41DB7F95 CRC64;

Query Match 91.9%; Score 440; DB 2; Length 395;
Best Local Similarity 95.0%; Pred. No. 7.6e-20;
Matches 95; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60
Db 213 LKEIDESDSEYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEKVEDFK 272

Qy 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98
Db 273 NSDGEQAQYLAAAEEDLIAKKAELKAEADLKAVDEPE 312

RESULT 5
Q9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RD Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RD Beall B.W.;
RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
```



OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=232;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=232;  
RX Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF255551; AAF68104.1; -;  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 1 1  
FT NON\_TER 262 262  
SQ SEQUENCE 262 AA; 32C769099466A584 CRC64;  
  
Query Match 90.0%; Score 431; DB 2; Length 262;  
Best Local Similarity 93.0%; Pred. No. 1.9e-19;  
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2  
  
Qy 1 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 60  
Db 65 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 124  
  
Qy 61 NSDGEQA-QYLAABEDL-AKKALEKTEADLKKAVHEPE 98  
Db 125 NSNGEQAQYRAAABEDLAAKQAELEKTEADLKKAVNEPE 164  
  
RESULT 8  
Q9L577 PRELIMINARY; PRT; 415 AA.  
ID Q9L577 PRELIMINARY; PRT; 415 AA.  
AC Q9L577;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE PspA (fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BG6692;  
RX MEDLINE=20448953; PubMed=10992499;  
RX DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of pspA; mosaic genes and evidence for past recombination  
RT in Streptococcus pneumoniae";  
RL EMBL; AF071808; AAF27704.1; -;  
DR InterPro; IPR000533; Tropomyosin.  
FT NON\_TER 415 415  
FT NON\_TER 45593 45593  
SQ SEQUENCE 415 AA; 41375ACBFA10PA46 CRC64;  
  
Query Match 90.0%; Score 431; DB 2; Length 415;  
Best Local Similarity 93.0%; Pred. No. 2.9e-19;  
Matches 93; Conservative 4; Mismatches 1; Indels 2; Gaps 2  
  
Qy 1 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 60  
Db 229 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 288  
  
Qy 61 NSDGEQA-QYLAABEDL-AKKALEKTEADLKKAVHEPE 98

DR EMBL; AF254258; AAF68093.1; -;  
DR InterPro; IPR009082; His\_kin\_homodim.  
FT NON\_TER 1 1  
FT NON\_TER 225 225  
SQ SEQUENCE 225 AA; F878A7618B72A692 CRC64;  
  
Query Match 90.2%; Score 432; DB 2; Length 225;  
Best Local Similarity 93.0%; Pred. No. 1.5e-19;  
Matches 93; Conservative 3; Mismatches 2; Indels 2; Gaps 2  
  
Qy 1 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 60  
Db 34 LKEIDESDSDYVKEGLRAPLQSELDKAKLSKLELSDKIDELDAETAKLEKVEDPK 93  
  
Qy 61 NSDGEQA-QYLAABEDL-AKKALEKTEADLKKAVHEPE 98  
Db 94 NSNGEQAQYRAAABEDLAAKQAELEKTEADLKKAVNEPE 133  
  
RESULT 6  
Q9L577 PRELIMINARY; PRT; 222 AA.  
ID Q9L577 PRELIMINARY; PRT; 222 AA.  
AC Q9L577;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PspA (fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669 (2000).  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multidrug-resistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."; J.

Db 289 NSNGEAOEYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 328

## RESULT 9

```

Q9LAY8
Q9LAY8 PRELIMINARY; PRT; 416 AA.
AC Q9LAY8;
AC Q9LAY8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment) .
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC Streptococcus.
NCBI_TaxID=1313;
[1]
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BG8838;
RC MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RX Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RT Infect. Immun. 68:5889-5900(2000) .
RL EMBL: AF071807; AAF27703.1; -.
DR InterPro: IPR000533; Tropomyosin.
DR PRINTS; P001194; TROPOMYOSIN.
FT NON TER 416 416
SQ SEQUENCE 416 AA; 45987 MW; 990C8858BC6B12C7 CRC64;

```

Query Match	88.5%	Score 424;	DB 2;	Length 416;
Best Local Similarity	92.0%	Pred. NO. 7.8e-19;		
Matches 92;	Conservative	4;	Mismatches 2;	Indels 2;
			Gaps	2;

Qy	1	LKEIDSDSDYYKVEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKDVDFPK	60
Db	229	LKEIDSDSDYYKVEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAETAKLEKDVDFPK	288

61 NSDGEQA-QYLA AAEEDL-AKKA ELEKTEADLKKAVHEPE 98

db 289 NSNGEOAEQYRAAGEDLAAKQAELEKTEADLKKA VNEPE 328

## RESULT 10

Account ID	Q9L5B5	PRELIMINARY;	PRT;	194 AA.
ID	Q9L5B5			
AC	Q9L5B5;			
AD	Q9L5B5;			
DT	01-OCT-2000 (T-EMBLrel. 15, Created)			
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)			
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)			
DE	PapA (Fragment).			
GN	Name=pspa;			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SP196;			
RX	MEDLINE=20472698; PubMed=11015380;			
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;			
RT	"pneumococcal pspa sequence types of prevalent multiresistant			
RT	pneumococcal strains in the United States and of internationally			
RT	disseminated clones.";			
RJ	J. Clin. Microbiol. 38:3663-3669(2000).			
RL				

Run	[Z]	Sequence from N.A.
RP		

RF SEQUENCE FROM  
RC STRAIN=SP196:

RA Beall B.W.;;

Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.  
EMBL: AF253407; AAF67355.1; -.

DR	InterPro:	IPR009082;	His_kin_homodim.
FT	NON_TER	1	1
FT	NON_TER	194	194
SQ	SEQUENCE	194 AA;	211116 MW; E68189FCA2B244F8 CRC64;

Query Match 88.3%; Score 423; DB 2; Length 194;  
Best Local Similarity 90.0%; Pred. No. 4.6e-19;

QY	1	LKEIDESDSDYVKEGLRAPIQSELDAKQAKLSKEELSDKIDELDAETIAKLEKDVEDFK	60
		:           :           :	
		:           :           :	
Db	55	LKEIDESDSEDYIKEGLRAPIQSKLDAKKAKLSKEELSDKIDELDAETIAKLEKDVEDFK	110

61 NSDGE0A-OYLAAAEEDL-AKKAELEKTEADLKAVHEPE 98

115 NSDGEQAEQYVAAKKDLDNAKKAELIENTEADLKKAQDEPE 154

## RESULT 11

Q6UEB2	PRELIMINARY;	PRT;	218 AA.
ID	Q6UEB2		
AC	Q6UEB2;		
DT	05-JUL-2004	(TREMBLrel. 27, Created)	
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)	
DE	PePa (Fragment).		
GN	Name=pspa,		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313;		
[1]	RN		
RP	SEQUENCE FROM N.A.		

RX PubMedID:14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;  
 RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;  
 RT "Epitope mapping of a protective monoclonal antibody against  
 RT Pneumocystis carinii with shared reactivity to Streptococcus  
 RL pneumoniae surface antigen pspA.";  
 RL Infect. Immun. 72:1548-1556(2004).

DR EMBL; AY31665; AAR20918.1; -.  
DR InterPro: IPR009082; His kin homodim.

DA	INTELLIGENCE, TERRORISM, AND ANTI-TERRO
PT	NON TER 1 1

SQ SEQUENCE Z18 AA; 22528 MW; 0F3FZ/EDECA06D/Z CAC04;

Query Match 88.3%; Score 423; DB 2; Length 218;  
Best Local Similarity 90.0%; Pred. No. 5.1e-19;  
Matches 90: Conservative 5; Mismatches 3; Indels 2; Gaps 2;

QY	1	LKEIDESDSYVKEGLRAPIQSELDAKQAKLSKLIELSDKIDELDAETAKLEKDVEDFK	60
		:           :           :	
Dh	27	LKEIDESDSYIKEGIRAPIOSKLDAAKKAKLSKLEISDKIDELDAETAKLEKDVEDFK	86

61 NSDGEOA-OYLAAAEEDL-AKKAEELEKTEADLKAVHEPE 98

db 87 NSDGE0AE0YI.VAAKK0I.DAKKAEI.ENTENDI.KKAYNEPE 126

## RESULT 12

Q9L568	PRELIMINARY;	PRT;	233 AA.
ID	Q9L568		
AC	Q9L568;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	pspA (fragment).		
GN	Name=pspA;		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313;		
FN	[1]		

```

QY 61 NSDGEQA-QYLAABEDL-AKKALEKTEADLKKAVHEPE 98
||||| ||| :||| ||||| ||||| |||
Db 109 NSDGEQAQYLVAAKKDLDAKKALENTADLKKAVDEPE 148

RESULT 14
Q9L564
ID Q9L564 PRELIMINARY; PRT; 243 AA.
AC Q9L564;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE pspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=152;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF255906; AAF70096.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1
FT NON TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 88.38; Score 423; DB 2; Length 243;
Best Local Similarity 90.04; Pred. No. 5.66-19;
Matches 90; Conservative 5; Mismatches 3; Indels 2; Gaps 2;

QY 1 LKEDISSDYVKEGLRAPLQSGELDAQAQKSLKEELSKIDELDAEIAKLEKDVDFK 60
||||| ||| :||| ||||| ||||| |||
Db 74 LKEDISSDYVKEGLRAPLQSGELDAQAQKSLKEELSKIDELDAEIAKLEKDVDFK 133

QY 61 NSDGEQA-QYLAABEDL-AKKALEKTEADLKKAVHEPE 98
||||| ||| :||| ||||| ||||| |||
Db 134 NSDGEQAQYLVAAKKDLDAKKALENTADLKKAVDEPE 173

RESULT 15
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE pspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=90;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally

```

```

RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=90;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255903; AAF70093.1;
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match      88.3%; Score 423; DB 2; Length 243;
Best Local Similarity 90.0%; Pred. No. 5.6e-19;
Matches 90; Conservative 5; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LKEIDSEDSYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKVEDFK 60
Db 50 LKEIDSEDSYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKVEDFK 109

Qy 61 NSDGEQA-QYLAABEDL-AKKAELKTEADLKAVHEPE 98
Db 110 NSDGEQAQYLVAAKQKLDKAKALENTEADLKAVDEPE 149

```

Search completed: June 18, 2005, 17:01:33  
Job time : 61.7417 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-2

Perfect score: 489

Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKKAHVEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482	98.6	550	ADK48356	Adk48356 Streptococ
2	482	98.6	550	ADR95223	Adr95223 Novel S.
3	429	87.7	194	AAW14584	AAW14584 Streptococ
4	429	87.7	194	ABW02618	ABW02618 Db16ac pn
5	429	87.7	8991	ABU08487	ABU08487 S. pneumo
6	424	86.7	183	AAW14570	AAW14570 Streptococ
7	424	86.7	183	ABW02604	ABW02604 Bg9739c p
8	422	86.3	168	ABW02609	ABW02609 L81905c p
9	406.5	83.1	167	AAW14575	AAW14575 Streptococ
10	395	80.8	166	AAW14568	AAW14568 Streptococ
11	395	80.8	166	ABW02602	ABW02602 Bg8743c p
12	385.5	78.8	185	AAW14566	AAW14566 Streptococ
13	385.5	78.8	185	ABW02600	ABW02600 Ac94c pne
14	352.5	72.1	204	AAW14571	AAW14571 Streptococ
15	352.5	72.1	204	ABW02605	ABW02605 Bf1019c p
16	336.5	68.8	198	ABW02615	ABW02615 Rxlc pneum
17	336.5	68.8	215	AAW04375	AAW04375 Streptococ
18	336.5	68.8	619	AAW63437	AAW63437 Pneumococ
19	336.5	68.8	619	AAW87598	AAW87598 Pneumococ
20	336.5	68.8	619	AAW86911	AAW86911 Pneumococ
21	336.5	68.8	619	AAW14838	AAW14838 Streptococ
22	336.5	68.8	619	AAW18782	AAW18782 S. pneumo
23	336.5	68.8	619	ABU45778	ABU45778 Protein e
24	336.5	68.8	619	AD052126	AD052126 Streptococ
25	336.5	68.8	648	AAW70336	AAW70336 Pneumococ

## ALIGNMENTS

## RESULT 1

ADK48356  
ID ADK48356 standard; protein; 550 AA.

AC ADK48356;

XX 20-MAY-2004 (first entry)

XX Streptococcus pneumoniae protein, Seq ID No 4871.

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

XX Streptococcus pneumoniae.

XX US6699703-B1.

XX 02-MAR-2004.

XX 26-MAY-2000; 2000US-00583110.

XX 02-JUL-1997; 97US-0051553P.

XX 12-MAY-1998; 98US-0085131P.

XX 30-JUN-1998; 98US-00107433.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;

XX WPI; 2004-212399/20.

XX N-PSDB; ADK45695.

XX New nucleic acid molecules and polypeptides useful for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, e.g. Streptococcus pneumoniae infection, and in drug screening.

XX Disclosure; SEQ ID NO 4871; 301pp; English.

XX The invention relates to isolated Streptococcus pneumoniae nucleic acids and polypeptides. The nucleic acids and proteins are useful for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, such as S. pneumoniae infection. These may also be used for drug screening procedures. The present sequence represents a Streptococcus pneumoniae polypeptide of the invention. Note: The sequence data for this patent did not appear in the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 550 AA;

Query Match 98.6%; Score 482; DB 8; Length 550;  
Best Local Similarity 99.0%; Pred. No. 5.8e-34;  
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEDISSDSYVKEGLRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 60  
Db 144 LKEDISSDSYVKEGLRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 203

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
Db 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 2

ADR95223 ID ADR95223 standard; protein; 550 AA.

XX AC ADR95223;

XX DT 16-DEC-2004 (first entry)

XX DE Novel S. pneumoniae protein sequence, SEQ ID 3858.

XX KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;

XX KW bacterial infection.

XX OS Streptococcus pneumoniae.

XX US US6800744-B1.

XX PD 05-OCT-2004.

XX PF 30-JUN-1998; 98US-00107433.

XX PR 02-JUL-1997; 97US-0051553P.

XX PR 12-MAY-1998; 98US-0085131P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI: 2004-697205/68.

XX DR N-PSDB; ADR92620.

XX New isolated nucleic acid encoding a Streptococcus pneumoniae polypeptide, useful for diagnosing, preventing and/or treating pathological conditions resulting from the bacterial infection.

XX PS Disclosure; SEQ ID NO 3858; 151pp; English.

XX The invention relates to an isolated nucleic acid comprising a sequence encoding a Streptococcus pneumoniae ADR91366polypeptide, or its fragments, with any of 9 fully defined sequences (appearing as ADR94308, ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682, ADR96079) or any of the fully defined sequences appearing as ADR91705, ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide sequences, or at least 40, 60 or 300 consecutive nucleotides, which is hybridisable under high stringency conditions to the nucleotide sequence. The nucleic acids and proteins are chosen from 5206 disclosed sequences. Also included are a recombinant expression vector comprising the isolated nucleic acid cited above operably linked to a transcription regulatory element, a cell comprising the recombinant expression vector and a probe comprising at least 20 consecutive nucleotides of the nucleotide sequences as cited above. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of pathological conditions resulting from bacterial infection by Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and otitis media. The present sequence is one of the 2603 disclosed S. pneumoniae protein sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in

CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.

XX SQ Sequence 550 AA;

Query Match 98.6%; Score 482; DB 8; Length 550;  
Best Local Similarity 99.0%; Pred. No. 5.8e-34;  
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEDISSDSYVKEGLRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 60  
Db 144 LKEDISSDSYVKEGLRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 203

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
Db 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 3

AAW14584 ID AAW14584 standard; protein; 194 AA.

XX AC AAW14584;

XX DT 17-OCT-2003 (revised)

XX DT 28-OCT-1997 (first entry)

XX DE Streptococcus pneumoniae PspA central region.

XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis; bacteraemia; pneumonia.

XX OS Streptococcus pneumoniae; strain Db16.

XX FH Key Location/Qualifiers

FT Misc-difference 61

FT /note= "unidentified amino acid"

XX W09709994-A1.

XX PD 20-MAR-1997.

XX PF 16-SEP-1996; 96WO-US014819.

XX PR 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI: 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain Db16. Comparison of the N-terminal and central regions (AAW14533-57 and AAW14562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in vaccines to protect animals against S. pneumoniae infection and hence for the prevention of diseases such as otitis media, meningitis, bacteraemia and pneumonia. The sequence of the 3' half of the PspA alpha-helical region and the immediate 5' tip of the coding sequence are likely to be the critical sequences for predicting PspA cross-reactions and vaccine composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 194 AA;

Query Match 87.7%; Score 429; DB 2; Length 194;  
 Best Local Similarity 89.9%; Pred. No. 7.8e-30;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEP 99  
 DB 61 XSDGEQAGQYLAAREEDLAAKQAELEKTEADLKKAVNEP 99

RESULT 4  
 ABW02618  
 ID ABW02618 standard; protein; 194 AA.  
 AC ABW02618;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Db16ac pneumococcal surface protein A (PspA) central region.  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1. .194  
 FT /note= "Xaa = Unknown amino acid"  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 DR WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for  
 detecting the presence of Streptococcus pneumoniae or its strain,  
 comprises at least two different full length isolated gene encoding  
 pneumococcal surface protein A.

Example 6; SEQ ID NO 64; 121pp; English.

The present invention relates to an immunological composition comprising  
 at least 2 different full length isolated genes encoding pneumococcal  
 surface protein A (PspAs) from different groups based on restriction  
 fragment polymorphism analysis. The invention is useful for obtaining  
 expression products by recombinant techniques to detect, determine,  
 isolate or diagnose the presence of Streptococcus pneumoniae or its  
 strain. The expression product is useful for preparing antigenic,  
 immunological or vaccine compositions, for eliciting antibodies, an  
 immunological response (other than or additional to antibodies), or a  
 protective response (including antibody or other immunological response  
 by administering compositions to a host). The invention is also useful as  
 vaccines and in gene therapy. The present sequence is Db16ac pneumococcal  
 surface protein A (PspA) central region. This sequence is used in the  
 exemplification of the invention

Query Match 87.7%; Score 429; DB 7; Length 194;  
 Best Local Similarity 89.9%; Pred. No. 6.1e-28;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 DB 7537 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 7596

Best Local Similarity 89.9%; Pred. No. 7.8e-30;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEP 99  
 DB 61 XSDGEQAGQYLAAREEDLAAKQAELEKTEADLKKAVNEP 99

RESULT 5  
 ABU08487  
 ID ABU08487 standard; protein; 8991 AA.  
 XX  
 AC ABU08487;  
 XX  
 DT 24-JUN-2003 (first entry)  
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1. .8991  
 FT /note= "All Xaa residues within this sequence are  
 unknown"  
 XX  
 PN US6500613-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 16-SEP-1996; 96US-00714741.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 XX (UYAL-) UNIV ALABAMA.  
 PA  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 2003-361534/34.

Isolated PspC amino acid sequence used as polymerase chain reaction or  
 hybridization probe, comprises pneumococcal surface protein having alpha-  
 helical, proline rich and repeat regions.

Disclosure; Col 145-188; 186pp; English.

The present invention relates to the isolation of Streptococcus  
 pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 like protein having alpha-helical, proline rich and repeat regions. The  
 PspC and PspA proteins may be used in a vaccine to protect against  
 pneumococcal infections. The polynucleotide sequences encoding PspC and  
 PspA may be used for the expression of the proteins, and as PCR primers  
 or hybridisation probes. The present sequence represents S. pneumoniae  
 PspA protein

Query Match 87.7%; Score 429; DB 6; Length 8991;  
 Best Local Similarity 89.9%; Pred. No. 6.1e-28;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
 DB 7537 LKEIDSDSDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 7596

```
QY      61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEP 99
Db      7597 XSDGEQAGQYLAABEDLIAKKAELEQTEADLKKAVNEP 7635

RESULT 6
AAW14570
ID      AAW14570 standard; protein; 183 AA.
XX
AC      AAW14570;
XX
XX      17-OCT-2003 (revised)
Dt      28-OCT-1997 (first entry)
XX
DE      Streptococcus pneumoniae PspA central region.
XX
KW      PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW      bacteraemia; pneumonia.
XX
OS      Streptococcus pneumoniae; strain Bg9739.
XX
PN      WO9709994-A1.
XX
PD      20-MAR-1997.
XX
PF      16-SEP-1996; 96WO-US014819.
XX
PR      15-SEP-1995; 95US-00529055.
XX
PA      (UABR-) UAB RES FOUND.
XX
XX      Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI      Hollingshead S, Tart R, Brooks-Walter A;
XX      WPI; 1997-202002/18.
XX
XX      Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT      in vaccines for protecting animals against S.pneumoniae infection.
XX
XX      Example 6; Fig 13; 296pp; English.
XX
XX      This sequence shows the central portion, including the C-terminus of the
CC      alpha-helix region and some of the proline-rich region, of pneumococcal
CC      surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.
CC      Comparison of the N-terminal and central regions (AAW14533-57 and
CC      AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC      be used to divide the strains into several families based on sequence
CC      homologies. PspA polypeptides, or fragments of them, can be used in
CC      vaccines to protect animals against S. pneumoniae infection and hence for
CC      the prevention of diseases such as otitis media, meningitis, bacteraemia
CC      and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC      region and the immediate 5' tip of the coding sequence are likely to be
CC      the critical sequences for predicting PspA cross-reactions and vaccine
CC      composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ      Sequence 183 AA;

Query Match      86.7%; Score 424; DB 2; Length 183;
Best Local Similarity 88.0%; Pred. No. 2e-29;
Matches 88; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAIAKLEKNVEDFK 60
Db      1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAIAKLEKNVEDFK 60
XX
QY      61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEP 100
Db      61 NSDGEQAGQYLAABEDLIAKKAELEKADLKKAVDEPE 100

RESULT 8
ABW02609
ID      ABW02609 standard; protein; 168 AA.
XX
AC      ABW02609;
XX
XX      12-FEB-2004 (first entry)
Dt
XX
```

ABW02604 standard; protein; 183 AA.

ABW02604;

12-FEB-2004 (first entry)

Bg9739c pneumococcal surface protein A (PspA) central region.

Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine; immunological; gene therapy; immunostimulant.

Unidentified.

US6592876-B1.

15-JUL-2003.

15-SEP-1995; 95US-00529055.

20-APR-1993; 93US-00048896.

06-JUN-1995; 95US-00465746.

(UABR-) UAB RES FOUND.

Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

Example 6; SEQ ID NO 50; 121pp; English.

The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is Bg9739c pneumococcal surface protein A (PspA) central region. This sequence is used in the exemplification of the invention

Sequence 183 AA;

Query Match 86.7%; Score 424; DB 7; Length 183;

Best Local Similarity 88.0%; Pred. No. 2e-29;

Matches 88; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAIAKLEKNVEDFK 60

Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEP 100

Db 61 NSDGEQAGQYLAABEDLIAKKAELEKADLKKAVDEPE 100

RESULT 8

ABW02609

ID ABW02609 standard; protein; 168 AA.

XX

AC ABW02609;

XX

12-FEB-2004 (first entry)

XX



DE L81905c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1.168  
 FT /note= "Xaa = Unknown amino acid"  
 XX  
 PN US6592876-B1.  
 XX  
 XX 15-JUL-2003.  
 PD  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 XX  
 DR  
 XX  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 55; 121pp; English.  
 PS  
 XX  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is L81905c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 XX Sequence 168 AA;  
 SQ  
 Query Match 86.3%; Score 422; DB 7; Length 168;  
 Best Local Similarity 88.0%; Pred. No. 2.7e-29;  
 Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60  
 QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 100  
 DB 61 NSDGEQAGQYLAAAEEDLTAQKALEKAFADLKKAHVEPE 100  
 RESULT 9  
 ID AAW14575  
 XX AAW14575 standard; protein; 167 AA.  
 XX  
 XX AAW14575;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain L81905.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 37  
 FT /note= "unidentified amino acid"  
 FT Misc-difference 41  
 FT /note= "unidentified amino acid"  
 FT Misc-difference 83  
 FT /note= "unidentified amino acid"  
 XX  
 XX WO9709994-A1.  
 PN  
 XX 20-MAR-1997.  
 PD  
 XX 16-SEP-1996; 96WO-US014819.  
 PF  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 PT  
 XX Example 6; Fig 13; 296pp; English.  
 PS  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain AAL81905.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 167 AA;  
 SQ  
 Query Match 83.1%; Score 406.5; DB 2; Length 167;  
 Best Local Similarity 87.0%; Pred. No. 6.2e-28;  
 Matches 87; Conservative 3; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLS-LEEKSDKIDELDAETAKLEKNVEDFK 59  
 QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 100  
 DB 60 NSDGEQAGQYLAAAEEDLTAQKALEKAFADLKKAHVEPE 99  
 RESULT 10  
 ID AAW14568  
 XX AAW14568 standard; protein; 166 AA.  
 XX  
 XX AAW14568;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE

```
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Bg8743.
OS
PN WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
DR
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 166 AA;
SQ
Query Match 80.8%; Score 395; DB 2; Length 166;
Best Local Similarity 82.0%; Pred. No. 6.3e-27;
Matches 82; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKNVEDFK 60
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEKXVDGDFP 60
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 61 NSNGEQAQYRAAEEDLAAKQAELEKTEADLKAVHEPE 100
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
DB 61 NSDGEQAGQYLVAEKDLDAKEAEELGNTGADLKAVDEPE 100
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
RESULT 11
ABW02602
ID ABW02602 standard; protein; 166 AA.
XX
XX ABW02602;
AC
XX 12-FEB-2004 (first entry)
DT
XX Bg8743c pneumococcal surface protein A (PspA) central region.
DE
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
OS
XX US6592876-B1.
PN
XX 15-JUL-2003.
PD
```

```
XX 15-SEP-1995; 95US-00529055.
PF
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 48; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspAs) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antigenic,
XX immunological or vaccine compositions, for eliciting antibodies, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Bg8743c
XX pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
XX Sequence 166 AA;
SQ
Query Match 80.8%; Score 395; DB 7; Length 166;
Best Local Similarity 82.0%; Pred. No. 6.3e-27;
Matches 82; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKNVEDFK 60
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEKXVDGDFP 60
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 61 NSNGEQAQYRAAEEDLAAKQAELEKTEADLKAVHEPE 100
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
DB 61 NSDGEQAGQYLVAEKDLDAKEAEELGNTGADLKAVDEPE 100
||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
RESULT 12
AAW14566
ID AAW14566 standard; protein; 185 AA.
XX
XX AAW14566;
AC
XX 17-OCT-2003 (revised)
DT
XX 28-OCT-1997 (first entry)
DT
XX Streptococcus pneumoniae PspA central region.
DE
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Ac94.
OS
XX WO9709994-A1.
PN
XX 20-MAR-1997.
PD
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
```

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 185 AA;  
 SQ  
 Query Match 78.8%; Score 385.5; DB 2; Length 185;  
 Best Local Similarity 81.2%; Pred. No. 4.8e-26;  
 Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAELEKTEADLKKAVHEPE 59  
 DB 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNKKQVDF 60  
 QY 60 KNSNGEQAFQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
 DB 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101  
 RESULT 14  
 AAW14571  
 ID AAW14571 standard; protein; 204 AA.  
 XX  
 AC AAW14571;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Bf1019.  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 185 AA;  
 SQ  
 Query Match 78.8%; Score 385.5; DB 2; Length 185;  
 Best Local Similarity 81.2%; Pred. No. 4.8e-26;  
 Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAELEKTEADLKKAVHEPE 59  
 DB 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNKKQVDF 60  
 QY 60 KNSNGEQAFQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
 DB 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101  
 RESULT 13  
 ABW02600  
 ID ABW02600 standard; protein; 185 AA.  
 XX  
 AC ABW02600;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX  
 DE Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 XX WPI; 2003-862841/80.  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,

PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 46; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 XX Sequence 185 AA;  
 SQ  
 Query Match 78.8%; Score 385.5; DB 7; Length 185;  
 Best Local Similarity 81.2%; Pred. No. 4.8e-26;  
 Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAELEKTEADLKKAVHEPE 59  
 DB 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNKKQVDF 60  
 QY 60 KNSNGEQAFQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
 DB 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101  
 RESULT 14  
 AAW14571  
 ID AAW14571 standard; protein; 204 AA.  
 XX  
 AC AAW14571;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Bf1019.  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Efi1019.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 204 AA;

Query Match 72.1%; Score 352.5; DB 2; Length 204;  
Best Local Similarity 74.0%; Pred. No. 4.2e-23;  
Matches 74; Conservative 12; Mismatches 13; Indels 1; Gaps 1;  
Qy 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60  
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEQDLKAAE 60  
Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEPE 100  
Db 61 ENNVVE-DYFKEGLEKTIAAKAELEKTEADLKAVNEPE 99

#### RESULT 15

ABW02605  
ID ABW02605 standard; protein; 204 AA.

XX AC ABW02605;

XX DT 12-FEB-2004 (first entry)

XX DE Efi1019c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX FN US5592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX DR WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 51; 121pp; English.

XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Efi1019c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX SQ Sequence 204 AA;

Query Match 72.1%; Score 352.5; DB 7; Length 204;  
Best Local Similarity 74.0%; Pred. No. 4.2e-23;  
Matches 74; Conservative 12; Mismatches 13; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60  
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEQDLKAAE 60  
Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEPE 100  
Db 61 ENNVVE-DYFKEGLEKTIAAKAELEKTEADLKAVNEPE 99

Search completed: June 18, 2005, 16:51:19  
Job time : 74.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-2  
Perfect score: 489  
Sequence: 1 LKEIDESDSEYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/protdata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/protdata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/protdata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/protdata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/protdata/1/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/protdata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	489	100.0	100	4	US-09-147-875A-2
2	482	98.6	550	4	US-09-583-110-4871
3	482	98.6	550	4	US-09-107-433-3858
4	479	98.0	100	4	US-09-147-875A-3
5	471.5	96.4	101	2	US-08-710-749-1
6	464.5	95.0	101	2	US-08-710-749-2
7	445	91.0	100	4	US-09-147-875A-5
8	438	89.6	98	4	US-09-147-875A-1
9	429	87.7	194	4	US-08-529-055-64
10	429	87.7	8931	4	US-08-714-741-32
11	428	87.5	100	4	US-09-147-875A-4
12	427.5	87.4	101	2	US-08-710-749-4
13	424	86.7	183	4	US-08-529-055-50
14	422	86.3	168	4	US-08-529-055-55
15	420.5	86.0	99	2	US-08-710-749-9
16	419	85.7	100	4	US-09-147-875A-6
17	417.5	85.4	101	2	US-08-710-749-3
18	408.5	83.5	101	2	US-08-710-749-5
19	406	83.0	100	4	US-09-147-875A-8
20	395	80.8	166	4	US-08-529-055-48
21	393.5	80.3	101	2	US-08-710-749-7
22	392.5	80.3	101	4	US-09-147-875A-9
23	385.5	78.8	185	4	US-08-529-055-46
24	380	77.7	100	4	US-09-147-875A-7
25	375	76.7	102	2	US-08-710-749-8
26	367.5	75.2	101	2	US-08-710-749-6
27	352.5	72.1	99	2	US-08-710-749-10

28	352.5	72.1	99	4	US-09-147-875A-11	Sequence 11, Appl
29	352.5	72.1	204	4	US-08-529-055-51	Sequence 51, Appl
30	346	70.8	100	4	US-09-147-875A-12	Sequence 12, Appl
31	336.5	68.8	99	2	US-08-710-749-11	Sequence 11, Appl
32	336.5	68.8	198	4	US-08-529-055-61	Sequence 61, Appl
33	336.5	68.8	619	1	US-08-465-746-2	Sequence 2, Appl
34	336.5	68.8	619	1	US-08-214-164-2	Sequence 2, Appl
35	336.5	68.8	619	2	US-08-467-852A-3	Sequence 3, Appl
36	336.5	68.8	619	2	US-08-246-636-2	Sequence 3, Appl
37	336.5	68.8	619	2	US-08-247-491A-3	Sequence 3, Appl
38	336.5	68.8	619	2	US-08-319-795-2	Sequence 2, Appl
39	336.5	68.8	619	2	US-08-468-985-2	Sequence 2, Appl
40	336.5	68.8	619	3	US-08-312-949-2	Sequence 2, Appl
41	336.5	68.8	648	1	US-08-072-070-2	Sequence 2, Appl
42	336.5	68.8	648	1	US-08-469-434-2	Sequence 2, Appl
43	336.5	68.8	648	1	US-08-214-223-2	Sequence 2, Appl
44	336.5	68.8	648	2	US-08-467-852A-2	Sequence 2, Appl
45	336.5	68.8	648	2	US-08-468-718-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-147-875A-2  
; Sequence 2, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-2

Query Match 100.0%; Score 489; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1e-38;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60  
DB 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAAEEDLAQKQAELEKTEADLKAVHEPE 100  
DB 61 NSNGEQAEQYRAAAEEDLAQKQAELEKTEADLKAVHEPE 100

RESULT 2  
US-09-583-110-4871  
; Sequence 4871, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4871

```

; LENGTH: 550
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4871

Query Match      98.6%; Score 482; DB 4; Length 550;
Best Local Similarity 99.0%; Pred. No. 3.5e-37;
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 144 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 203

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 3
US-09-107-433-3858
; Sequence 3858, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:

US-09-107-433-3858

Query Match      98.6%; Score 482; DB 4; Length 550;
Best Local Similarity 99.0%; Pred. No. 3.5e-37;
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 144 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 203

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 4
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

Query Match      98.0%; Score 479; DB 4; Length 100;
Best Local Similarity 98.0%; Pred. No. 8.6e-38;
Matches 98; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 5
US-08-710-749-1
; Sequence 1, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:

```



```

; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0; Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-64

Query Match 87.7%; Score 429; DB 4; Length 194;
Best Local Similarity 89.9%; Pred. No. 9e-33;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEP 99
Db 61 XSDGEQAGQYLAABEEDLIAKAELEQTEADLKAVNEP 99

RESULT 10
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0; Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match 87.7%; Score 429; DB 4; Length 8991;
Best Local Similarity 89.9%; Pred. No. 9.1e-31;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60
Db 7537 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 7596

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEP 99
Db 7597 XSDGEQAGQYLAABEEDLIAKAELEQTEADLKAVNEP 7635

RESULT 11
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-4

Query Match 87.5%; Score 428; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 5e-33;
Matches 90; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDYKGLRAPLQSELDKAKLSKLELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEP 100
Db 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEP 100
```



Db 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVIDEPE 100

RESULT 12  
US-08-710-749-4  
; Sequence 4, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-4

Query Match 87.4%; Score 427.5; DB 2; Length 101;  
Best Local Similarity 90.1%; Pred. No. 5.7e-33;  
Matches 91; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 LKEDISSDSDYVKEGLRAPLQSEL-DAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDF 59  
|||||  
Db 1 LKEDISSDSDYVKEGERAPLQSELDADAKQAKLSKLEELSDKIDELDAEIAKLEKVD 60  
|||||

QY 60 KNSGEQAGQYRAAEEEDLAQKAELEKTEADLKKAVIDEPE 100  
|||||  
Db 61 KNSDGEQAGQYLAAGEDLIAKKAELKAELEQTEADLKKAVIDEPE 101  
|||||

RESULT 13  
US-08-529-055-50  
; Sequence 50, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 183 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-50

Query Match 86.7%; Score 424; DB 4; Length 183;  
Best Local Similarity 88.0%; Pred. No. 2.5e-32;  
Matches 88; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEDISSDSDYVKEGLRAPLQSELDADAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDF 60  
|||||  
Db 1 LKEDISSDSDYVREGFRAPLQSELDADAKQAKLSKLEELSDKIDELDAEIAKLEKVD 60  
|||||

QY 61 NSNGEQAGQYRAAEEEDLAQKAELEKTEADLKKAVIDEPE 100  
|||||  
Db 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVIDEPE 100  
|||||

RESULT 14  
US-08-529-055-55  
; Sequence 55, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-55

Query Match      86.3%; Score 422; DB 4; Length 168;
Best Local Similarity 88.0%; Pred. No. 3.4e-32;
Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKNVDFK 60
Db 1 LKEIDESSEDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKNVDFK 60

Qy 61 NSNGEQARQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 100
Db 61 NSDGEQACQYLAARAEEDLAAKKALEKAEADLKKAHVEPE 100

RESULT 15
US-08-710-749-9
; Sequence 9, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: amino acid
US-08-710-749-9

Query Match      86.0%; Score 420.5; DB 2; Length 99;
Best Local Similarity 92.1%; Pred. No. 2.5e-32;
Matches 93; Conservative 4; Mismatches 1; Indels 3; Gaps 3;

Qy 1 LKEIDESSEDYVKEGLRAPLQSEL-DAKQAKLSKLEELSDKIDELDAEIAKLEKNVDF 59
Db 1 LKEIDESSEDYVKEGLRAPLQSELDDAKQAKLSKLEELSDKIDELDAEIAKLEKVEDF 60

Qy 60 KNSNGEQARQYRAAAEEDLAAKQAELEKTEADLKKAHVEPE 100
Db 61 KNSDGEQARQYLAARAEEDLAAKKALEKTEADLKKAHVEPE 99

Search completed: June 18, 2005, 17:07:04
Job time : 18.9189 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-2  
Perfect score: 489  
Sequence: 1 LKEDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	489	100.0	100	15	US-10-674-755-2
2	479	98.0	100	15	US-10-674-755-3
3	445	91.0	100	15	US-10-674-755-5
4	438	89.6	98	15	US-10-674-755-1
5	429	87.7	194	15	US-10-299-636-79
6	428	87.5	100	15	US-10-674-755-4
7	424	86.7	183	15	US-10-299-636-65
8	422	86.3	168	15	US-10-299-636-70
9	419	85.7	100	15	US-10-674-755-6
10	406	83.0	100	15	US-10-674-755-8
11	395	80.8	166	15	US-10-299-636-63
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 79, Appli
					Sequence 4, Appli
					Sequence 65, Appli
					Sequence 70, Appli
					Sequence 6, Appli
					Sequence 8, Appli
					Sequence 63, Appli

12	392.5	80.3	101	15	US-10-674-755-9	Sequence 9, Appli
13	385.5	78.8	185	15	US-10-299-636-61	Sequence 61, Appli
14	380	77.7	100	15	US-10-674-755-7	Sequence 7, Appli
15	352.5	72.1	99	15	US-10-674-755-11	Sequence 11, Appli
16	352.5	72.1	204	15	US-10-299-636-66	Sequence 66, Appli
17	346	70.8	100	15	US-10-674-755-12	Sequence 12, Appli
18	336.5	68.8	198	15	US-10-299-636-76	Sequence 76, Appli
19	336.5	68.8	354	15	US-10-299-636-105	Sequence 105, Appli
20	336.5	68.8	588	15	US-10-299-636-96	Sequence 96, Appli
21	336.5	68.8	619	10	US-09-882-774-1	Sequence 1, Appli
22	336.5	68.8	619	15	US-10-282-122A-73702	Sequence 73702, A
23	336.5	68.8	619	16	US-10-414-532-72	Sequence 72, Appli
24	335.5	68.6	170	15	US-10-299-636-75	Sequence 75, Appli
25	335.5	68.6	181	15	US-10-299-636-57	Sequence 57, Appli
26	335.5	68.6	643	15	US-10-299-636-95	Sequence 95, Appli
27	335.5	68.6	670	9	US-09-748-875-63	Sequence 63, Appli
28	335.5	68.6	670	10	US-09-298-523B-63	Sequence 63, Appli
29	335.5	68.6	690	9	US-09-748-875-61	Sequence 61, Appli
30	335.5	68.6	690	10	US-09-298-523B-61	Sequence 61, Appli
31	335.5	68.6	691	9	US-09-748-875-1	Sequence 1, Appli
32	335.5	68.6	691	10	US-09-298-523B-1	Sequence 1, Appli
33	335.5	68.6	701	9	US-09-748-875-62	Sequence 62, Appli
34	335.5	68.6	701	10	US-09-298-523B-62	Sequence 62, Appli
35	335.5	68.6	707	9	US-09-748-875-2	Sequence 2, Appli
36	335.5	68.6	707	10	US-09-298-523B-2	Sequence 2, Appli
37	335.5	68.6	711	9	US-09-748-875-3	Sequence 3, Appli
38	335.5	68.6	711	10	US-09-298-523B-3	Sequence 3, Appli
39	335.5	68.6	739	17	US-10-732-923-3294	Sequence 3294, Ap
40	335.5	68.6	929	9	US-09-748-875-60	Sequence 60, Appli
41	335.5	68.6	929	10	US-09-298-523B-60	Sequence 60, Appli
42	335.5	68.6	929	15	US-10-299-636-94	Sequence 94, Appli
43	332.5	68.0	188	15	US-10-299-636-74	Sequence 74, Appli
44	328.5	67.2	99	15	US-10-674-755-16	Sequence 16, Appli
45	327.5	67.0	99	15	US-10-674-755-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1  
US-10-674-755-2  
; Sequence 2, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-2

Query Match 100.0%; Score 489; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.1e-32;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEDSESDYVKEGLRAPQLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
DB 1 LKEDSESDYVKEGLRAPQLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
QY 61 NSNGEQAYRAAEEDLAAKQAELEKTEADLKAVHEPE 100  
DB 61 NSNGEQAYRAAEEDLAAKQAELEKTEADLKAVHEPE 100  
RESULT 2

```
US-10-674-755-3
; Sequence 3, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      98.0%; Score 479; DB 15; Length 100;
Best Local Similarity 98.0%; Pred. No. 2.6e-31;
Matches 98; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
DB 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 3
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match      91.0%; Score 445; DB 15; Length 100;
Best Local Similarity 92.0%; Pred. No. 1.4e-28;
Matches 92; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
DB 61 NSDGEQAGQYLAABEEDLIAKAELEQTEADLKKAVHEPE 100

RESULT 4
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
```

```
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match      89.6%; Score 438; DB 15; Length 98;
Best Local Similarity 94.0%; Pred. No. 5.1e-28;
Matches 94; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
DB 61 NSDGEQA-QYLAABEEDL-AKAELEKTEADLKKAVHEPE 98

RESULT 5
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match      87.7%; Score 429; DB 15; Length 194;
Best Local Similarity 89.9%; Pred. No. 5.7e-27;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 99
DB 61 XSDGEQAGQYLAABEEDLIAKAELEQTEADLKKAVNEP 99

RESULT 6
US-10-674-755-4
```

```
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match      87.5%; Score 428; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 3.3e-27;
Matches 90; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAGQYLAAGEDLLAKAELEKAEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match      86.7%; Score 424; DB 15; Length 183;
Best Local Similarity 88.0%; Pred. No. 1.4e-26;
Matches 88; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYVREGFRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAGQYLAAGEDLLAKAELEKAEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-70

Query Match      86.3%; Score 422; DB 15; Length 168;
Best Local Similarity 88.0%; Pred. No. 1.8e-26;
Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAGQYLAAGEDLLAKAELEKAEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-10-299-636-70
; Sequence 70, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa at position 38 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa at position 42 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70

Query Match      86.3%; Score 422; DB 15; Length 168;
Best Local Similarity 88.0%; Pred. No. 1.8e-26;
Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEQAGQYLAAGEDLLAKAELEKAEADLKKAHVPE 100
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-10-674-755-6
; Sequence 6, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match      85.7%; Score 419; DB 15; Length 100;
Best Local Similarity 88.0%; Pred. No. 1.8e-26;
Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVDEPE 100

RESULT 10
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-8

Query Match      83.0%; Score 406; DB 15; Length 100;
Best Local Similarity 85.0%; Pred. No. 1.9e-25;
Matches 85; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKGIIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIPKLEKNVEYFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100

RESULT 11
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111

; NAME/KEY: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      80.8%; Score 395; DB 15; Length 166;
Best Local Similarity 82.0%; Pred. No. 2.6e-24;
Matches 82; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEQAGQYLAAAEKDLDAKAEELGNTGADLKKAVDEPE 100

RESULT 12
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      80.3%; Score 392.5; DB 15; Length 101;
Best Local Similarity 82.2%; Pred. No. 2.4e-24;
Matches 83; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 59
Db 1 LKEIDSESDYVKEGLRVPLQSELDVQKAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 60 KNSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101

RESULT 13
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
```

; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-61

Query Match 78.8%; Score 385.5; DB 15; Length 185;  
Best Local Similarity 81.2%; Pred. No. 1.7e-23;  
Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAETAKLEKNVEDF 59  
DB 1 LKEIDSDSDYVKEGLRVLQSELDVQAKLLKLELSKIDELDAETAKLEKNVEDF 60  
QY 60 KNSGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
DB 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101

RESULT 14  
US-10-674-755-7  
; Sequence 7, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-7

Query Match 77.7%; Score 380; DB 15; Length 100;  
Best Local Similarity 80.0%; Pred. No. 2.4e-23;  
Matches 80; Conservative 7; Mismatches 13; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAETAKLEKNVEDF 60  
DB 1 LKEIDSDSDYVKEGLRAPLQSKLDAAKAKLSKLELSKIDEXSKKIDELDAETAKLEKDVDP 60  
QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
DB 61 NSDGEQAGQYLVAAEKDLDAKAEELGNTGADLKKAVDEPE 100

RESULT 15  
US-10-674-755-11  
; Sequence 11, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-11

Query Match 72.1%; Score 352.5; DB 15; Length 99;  
Best Local Similarity 74.0%; Pred. No. 3.9e-21;  
Matches 74; Conservative 12; Mismatches 13; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAETAKLEKNVEDF 60  
DB 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSKIDELDAETAKLEKNVEDF 60  
QY 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
DB 61 ENNVVE-DYFKSGLEKTIAAKKALEKTEADLKKAVNEPE 99

Search completed: June 18, 2005, 18:00:20  
Job time : 62.963 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-2  
Perfect score: 489  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	336.5	68.8	619	2 A97887	surface protein ps
2	336.5	68.8	619	2 A41971	surface protein ps
3	136	27.8	744	2 F95013	pneumococcal surfa
4	110.5	22.6	562	2 G70002	hypothetical prote
5	110	22.5	1169	2 A64505	Pil5 homolog - Met
6	110	22.5	1269	2 F84730	probable myosin he
7	107	21.9	1006	2 C70445	ATPase subunit of
8	105	21.5	522	2 G02533	occludin - human
9	105	21.5	3488	2 T34418	hypothetical prote
10	103.5	21.2	764	2 T05409	hypothetical prote
11	103.5	21.2	924	2 S06117	myosin heavy chain
12	103.5	21.2	2007	1 B43402	myosin heavy chain
13	102.5	21.0	886	2 H69378	conserved hypotet
14	102.5	21.0	896	2 S43074	epidermal growth f
15	102.5	21.0	4558	2 C82199	RTX toxin RtxA vcl
16	101.5	20.8	1376	2 A52522	myosin heavy chain
17	101	20.7	1138	2 T24635	hypothetical prote
18	99.5	20.3	161	2 S48396	tropomyosin TPM2 -
19	99.5	20.3	387	2 S57834	fcrA protein precu
20	99.5	20.3	1790	2 S67593	transport protein
21	99	20.2	281	2 F75216	hypothetical prote
22	99	20.2	1319	2 A28313	glued protein - fr
23	98.5	20.1	157	2 A97703	ATP synthase B cha
24	98.5	20.1	388	2 A46173	Mrp4 protein - Str
25	98.5	20.1	388	2 S52536	fcrA 15 protein -
26	98.5	20.1	405	2 A33939	Fc gamma (IgG) rec
27	98.5	20.1	1053	2 A41642	dynactin - chicken
28	98	20.0	1190	2 E84193	chromosome segrega
29	97	19.8	1156	2 B70356	chromosome assembl

30	96.5	19.7	201	2 A45332	zipper protein - c
31	96.5	19.7	622	2 G96703	unknown protein, 3
32	96.5	19.7	741	2 A26572	bsg25D protein - f
33	96.5	19.7	848	2 A44972	paramyosin - nemat
34	96.5	19.7	866	2 S04027	paramyosin - Caeno
35	96.5	19.7	872	2 T13296	hypothetical prote
36	96.5	19.7	1992	2 A47297	myosin heavy chain
37	96	19.6	407	1 EDBEQ3	immediate-early pr
38	96	19.6	1177	2 B75150	chromosome segrega
39	96	19.6	1959	1 A33977	myosin heavy chain
40	95.5	19.5	408	2 S30283	protein M precursor
41	95.5	19.5	415	2 S35760	fcrA protein precu
42	95.5	19.5	1110	2 I51116	NF-180 - sea lamp
43	95.5	19.5	1957	2 T38077	hypothetical coile
44	95	19.4	564	2 A60115	M protein precursor
45	95	19.4	1291	2 T09273	probable tail-host

ALIGNMENTS

RESULT 1

A97887  
surface protein psps precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:G  
C:Genetics:  
A:Gene: psps

Query Match

68.8%; Score 336.5; DB 2; Length 619;  
Best Local Similarity 70.0%; Pred. No. 3.6e-15;  
Matches 70; Conservative 15; Mismatches 14; Indels 1; Gaps 1;

QY 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60

DB 223 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 282

QY 61 NSNGQAEQYRAAEEDLAQKAELEKTEADLKAVHEPE 100

DB 283 ENNVVE-DYFKEGLEKTTAAKAELEKTEADLKAVHEPE 321

RESULT 2

A41971  
surface protein psps precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspsA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

[illegible]

```

Query Match      27.8%; Score 136; DB 2; Length 744;
Best Local Similarity 34.1%; Pred. No. 0.063;
Matches 45; Conservative 20; Mismatches 27; Indels 40; Gaps 6;

Qy      2 KEIDE-----SSESYVVEGRLRAPQSELDAAKQAKLSK-----LELSDKI-----D 43
      ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      314 KEISNLEILLGGPEDDT-----AALQNKLAAKKAELAKKQTELEKLLDSLPGEKQTQD 368

Qy      44 ELD--AIIAKLEKVEFDKNS-----NGEQAEQYRAAEEDLAAKQAELE 86
      ||| ||| : ||| : ||| : ||| : ||| : ||| :
Db      369 ELDKAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLATKKAELE 428

Qy      87 KTEADLKKAVHE 98
      ||| : ||| : ||| : ||| : ||| : ||| :
Db      429 KTOKELDAAALNE 440

```

hypothetical protein ytwp - *Bacillus subtilis*  
C:Species: *Bacillus subtilis*  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: G70002  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertone  
C.; Bron, S.; Brulet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabel, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
Y, M.; Ogawa, K.; Ogita, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
A:Authors: Schlecht, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tognoni, A.; Toso, V.; Uchiyama, K.  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; PMID:98044033; PMID:9384377  
A:Accession: G70002  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-562 <KUN>  
A:Cross-references: UNIPROT:O34894; GB:Z99119; GB:AL009126; NID:G2635411; PIDN:CAB14939  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ytwp

```

Query match. 22.6%; Score 110.5; DB 2; Length 302;
Best local similarity 32.4%; Pred. No. 2.3;
Matches 36; Conservative 21; Mismatches 33; Indels 21; Gaps 5;

Qy 2 KEIDSESDYVKEGLRAPIQSELDAAKQAKLKEE-----LSDKIDELDA----- 47
Db 325 KEHTKAETE-LVIESYTL-TAGELGKQQAPEKELDRIGKLLSVKDKLDAEHVAYSILVE 382

Qy 48 EIAKLEKNVEDFKNSNGEQAEQVRAAEEDLAAKQAELEKTEADLKKAVHE 98
Db 383 EVASIEKQIEEVKKEHAEYRENILQARKEBQARE-----TSLNKKTTTISE 428

```

RESULT 5  
A64505  
P115 homolog - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Jun-2000  
C:Accession: A64505  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID: 96337999; PMID: 8698087  
A:Accession: A64505  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1169 <BUL>  
A:Cross-references: GB:U67604; GB:L77117; NID:g1592224; PID:g1500543; TIGR:MJ1643  
C:Genetics:  
A:Map position: FOR1623481-1626990  
A:Superfamily: chromosome segregation protein SWC1

```

Qy      1  LKEIDE SDSDYVYKEGLRAPLQSELD-----AKQAKLSKLEELSDKIDELDAFIKLEKN 55
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      Best local similarity 31.2%; Pred. NO. 5;
      Query match 100%; Pos. 27; Length 128;
      Matches 35; Conservative 22; Mismatches 41; Indels 14; Gaps 3;

```

```

799 LKRNWETEGELKILEKAKLKNKNEIDKGLTFLVKELIPKLELNKKVSELINKKVILEKN 858
Oy      56 VEDFKNS---NGEQAEQYRAAAEE-----DLAAQAELEKTEADLKKAHVE 98
          :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db      859 ISPVKESIEKNLSILEBKRYRYEELAKNKLKELTEKKQLEKSETTLERERRE 910

RESULT 6
F84730
C:Probable myosin heavy chain [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84730
R:lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84730
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1269 <STO>
A:Cross-references: GB:AE002093; NID:G6598483; PIDN:AAC69932.2; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32240
A:Map position: 2

```

```

Query Match      22.5%; Score 110; DB 2; Length 1269;
Best Local Similarity 33.7%; Pred. No. 5.4;
Matches 34; Conservative 17; Mismatches 36; Indels 14; Gaps 3;

Qy      6   ESDSEYVVKGLRAPLOSELDAKOAKLSKLE-----ELSKIDELDARIKLEK 54
         |||::|||:|||:|||:|||::|||::|||
Db      661  EADSKGYL--GQVAELQSTLEAFQVKSSSLEAAALNIATENEKETLENLNAVTSKKKLEA 718

Qy      55  NVEDFKNSNGEQAEQYRAAAEEDLAQAQAELEKTEADLKKA 95
         ||::|||::|||::|||::|||::|||
Db      719  -TWDEYGVKTISE--SENLEGI RNELNVTQGKLEIENDLKAA 758

```

RESULT 7  
C70445  
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: C70445  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: C70445  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1006 <AOF>  
A:Cross-references: UNIPROT:O67588; GB:AE000750; NID:G2983999; PIDN:AA07550.1; PID:G298  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: clpB  
C:Superfamily: endopeptidase Clp ATP-binding chain  
C:Keywords: hydrolase

Query Match	21.9%;	Score 107;	DB 2;	Length 1006;
Best Local Similarity	32.7%;	Pred. No. 6.8;		
Matches	34; Conservative	22; Mismatches	32; Indels	16; Gaps 4;
QY	1 LKEIDSDSE----	DYVKEGLRAPQLSELDAKQAKLSK-LRELSKDIDELDAETIAKLEK	54	
	: : : :	: : : :   : :   : : :   : :   : :   : :		
Db	552 IKALEEQTITANLKGDIYEK-----	AQKIISKAKLEKEKQELLGKVGVGEATIAELKK	604	
QY	55 NVEDDFKNSNGEQAYQRAAAEEDLAARQAQAELEKTEADLKCAVHE	98		
	: : : :	: : : :   : :   : :   : :		

Db 605 KIEELDEKIEAAEKGDYKAEALKEKAKLEK---ELKKEQ 645

RESULT 8  
G02533  
occludin - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C:Accession: G02533  
R:Van Itallie, C.M.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: H01403  
A:Accession: G02533  
A:Status: preliminary; translated from GB/EMBL/DDDBJ  
A:Molecule type: mRNA  
A:Residues: 1-522 <VAN>  
A:Cross-references: UNIPROT:Q16625; EMBL:U53823; NID:g1322281; PIDN:AAB00195.1;  
C:Superfamily: occludin

```
Query Match      21.5%; Score 105; DB 2; Length 522;  
Best Local Similarity 30.5%; Pred. No. 4.8;  
Matches 29; Conservative 25; Mismatches 25; Indels 16; Gaps 5;  
  
Qy    5 DESDSYVKEGLRAPLQSELDKQAK-----LSKLEELSDKIDELDAIAKLNVED 58  
||| : ||| : | : | : | : | : | : | : | : | : | :  
Db   410 DELE-EDWIRE--YPITSDQQRLVKRPNFTGLQEYKSLQSELDEINKELSRDLKELDD 466  
  
Qy   59 FKMSNGEQARQYPAABEDLAQAQELEKTEADLK 93  
:: : | : | : | : | : | : | : | : | : | :  
Db  467 YR----EESEYNAADAENRLKOV---KGSADYK 494
```

RESULT 9  
T34418  
hypothetical protein F12F3.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999. #text\_change 29-Oct-1999  
C;Accession: T34418  
R;Fulton, B.; Wohlmann, P.  
submitted to the EMBL Data Library, July 1998  
A;Description: The sequence of C. elegans cosmid F12F3.  
A;Reference number: Z21521  
A;Accession: T34418  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-3488 <FUL>  
A;Cross-references: EMBL:U80022; PIDN:AC25885.1; GSPDB:GN00023; CESP:F12F3.3  
A;Experimental source: strain Bristol N2; clone F12F3  
C;Genetics:  
A;Gene: CESP:F12F3.3  
A;Map position: 5  
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match	21.5%;	Score 105;	DB 2;	Length 3488;
Best Local Similarity	35.1%;	Pred. No. 31;		
Matches	40;	Conservative 19;	Mismatches 33;	Indels 22; Gaps 6;
Qy	2	KEIDES---DSBDYKVGURAPLQSELDQAQKLSKL-----EEISDKTIDELDAEI	49	
Db	1009	KETDEKLKDAETAATKQEADEKSLDA-QEKIKKVSDDAARKEKELNDKL-KLESEI	1066	
Qy	50	AKLEKNVDFKNSNGBQASQYRAAABEDLAQAQAELEK-----TEADLKAAVHE	98	
Db	1067	ATKKAASADKLKLV---EEOACAKAAABEAAKKOKEKDEOLKLDTEAAKSKAAAE	1117	

RESULT 10  
T05409  
hypothetical protein F10M6.170 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05409  
R:Sevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herr

submitted to the Protein Sequence Database, February 1998

A;Reference number: Z15414  
A;Accession: T05409  
A;Molecule type: DNA  
A;Residues: 1-764 <BEV>  
A;Cross-references: UNIPROT:O49371; EMBL:AL021811  
A;Experimental source: cultivar Columbia; BAC clone F10M6  
C;Genetics:  
A;Map position: 4  
A;Note: F10M6.170

Query Match 21.2%; Score 103.5; DB 2; Length 764;  
Best Local Similarity 32.5%; Pred. No. 8.8;  
Matches 37; Conservative 17; Mismatches 43; Indels 17; Gaps 3;

Qy 2 KEIDSESDYVKEGLRAPLQSELDKQAKLSKLE-----EELSDKIDELDAEIAKLE 53  
Db REIEELKHKLREDERAALQSLTLKEEELKMKRQETANRSKEVSMASFEFSKSLLS 222  
Qy 54 KNVEDFNKNSGEQAEQYRAAAEED-----LAAKQAELEK---TEADLKKAHVE 98  
Db 223 KANEVVKRQGEIYALQALEEKEBELEISKATKKLEQEKLETEANLKKQTEE 276

#### RESULT 11

S06117  
myosin heavy chain, nonmuscle (clone lambda-FMHC) - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
A;Accession: S06117  
R;Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.  
Eur. J. Biochem. 184, 611-616, 1989  
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in S.

A;Reference number: S06116; MUID:90032648; PMID:2806244

A;Accession: S06117  
A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-924 <KAT>

A;Cross-references: UNIPROT:Q02015; GB:X17590  
A;Note: This translation is not annotated in GenBank entry GGMHCFPC, release 114  
C;Superfamily: myosin heavy chain; myosin motor domain homology  
F:1-303/Domain: myosin motor domain homology (fragment) <MMOT>

Query Match 21.2%; Score 103.5; DB 2; Length 924;  
Best Local Similarity 28.1%; Pred. No. 11;  
Matches 36; Conservative 24; Mismatches 29; Indels 39; Gaps 5;

Qy 5 DESDSEYVKEGLRAPLQSELDKQAKLSKLE-----LSDK-----IDELD 46  
Db 595 ETTDLQDQIAE-----LQAQIEELKIQAKKEELQALARGDEAVQKNNALKVIRELQ 649

Qy 47 AEIAKLEKNVEDFNKNSGEQAEQYRAAAE-----DLAAKQAEI-----EKTEA 90  
Db 650 AQIAELQEDLSEKASRNKAEKQKRDLSLEALKTELEDLTDTTAAQQLRTKREQEVA 709

Qy 91 DLKKAHVE 98

Db 710 ELKKAIEE 717

#### RESULT 12

B43402

myosin heavy chain-B, neuronal - chicken

N;Contains: myosin ATPase (EC 3.6.4.1)

C;Species: Gallus gallus (chicken)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C;Accession: B43402; A43402

R;Takahashi, M.; Kawamoto, S.; Adelstein, R.S.

J. Biol. Chem. 267, 17864-17871, 1992

A;Title: Evidence for inserted sequences in the head region of nonmuscle myosin specific

A;Reference number: A43402; MUID:92388144; PMID:1355479

A;Accession: B43402  
A;Molecule type: mRNA  
A;Residues: 1-2007 <TAK>  
A;Cross-references: UNIPROT:Q02015; GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212452  
A;Note: the sequence of residues 212-221 and 632-652 and the corresponding nucleotide se

A;Accession: A43402  
A;Molecule type: mRNA  
A;Residues: 1-211,222-631;653-2007 <TA2>  
A;Cross-references: GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212449

A;Note: sequence extracted from NCBI backbone (NCBIN:112864)

C;Comment: Alternatively spliced segments 1 and 2 are found exclusively in nonmuscle myo  
C;Superfamily: myosin heavy chain; myosin motor domain homology  
C;Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated

F:1-2007/Product: myosin heavy chain-B, neuronal #status predicted <MYN>  
F:1-211,222-631,653-2007/Product: myosin heavy chain-B, neuronal #status predicted <MYN>  
F:88-802/Domain: myosin motor domain homology <MMOT>

F:178-185/Region: nucleotide-binding motif A (P-loop)

F:212-221/Region: alternatively spliced segment 1 #status experimental

F:559-593/Region: actin binding #status predicted

F:632-652/Region: alternatively spliced segment 2 #status experimental

F:692-714/Region: actin binding #status predicted

F:875-2007/Domain: coiled coil #status predicted <COI>

F:875-1315/Region: S2

F:1316-2007/Region: light meromyosin

F:129/Modified site: N6,N6-trimethyllysine (Lys) #status predicted

F:184/Binding site: ATP (Lys) #status predicted

F:732,742/Active site: Cys #status predicted

F:1954/Binding site: phosphate (Thr) (covalent) #status predicted

F:1987/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 21.2%; Score 103.5; DB 1; Length 2007;

Best Local Similarity 28.1%; Pred. No. 23;

Matches 36; Conservative 24; Mismatches 29; Indels 39; Gaps 5;

Qy 5 DESDSEYVKEGLRAPLQSELDKQAKLSKLE-----LSDK-----IDELD 46

Db 1094 ETTDLQDQIAE-----LQAQIEELKIQAKKEELQALARGDEAVQKNNALKVIRELQ 1148

Qy 47 AEIAKLEKNVEDFNKNSGEQAEQYRAAAE-----DLAAKQAEI-----EKTEA 90

Db 1149 AQIAELQEDLSEKASRNKAEKQKRDLSLEALKTELEDLTDTTAAQQLRTKREQEVA 1208

Qy 91 DLKKAHVE 98

Db 1209 ELKKAIEE 1216

#### RESULT 13

H69378

conserved hypothetical protein AF1032 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69378

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB9021

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 21.0%; Score 102.5; DB 2; Length 886;

Best Local Similarity 26.7%; Pred. No. 12;

Matches 35; Conservative 28; Mismatches 31; Indels 37; Gaps 6;

Qy 1 LKEIDSESDSEYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKN----- 55

```
A;Gene: VC1451
A;Map position: 1

Query Match          21.0%; Score 102.5; DB 2; Length 4558;
Best Local Similarity 38.3%; Pred. No. 59;
Matches 31; Conservative 11; Mismatches 26; Indels 13; Gaps 3;

QY      16 GLRAPIQLSEI-DAKQAKLISK-----BELSDIKDELDAETAKLKNVEDFKNSNGEQAEOY 70
        |||..|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
Db       1726 GLLDGVQSQDLDADKALANDKIARAANKOTLSDNNNSKVYESVAKSEAGV-----AQGEQN 1777

QY      71 RAAAEEDLAQAELKTEAD 91
        |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
Db       1778 RAGVEQDIADAQADAERKAD 1798

Search completed: June 18, 2005, 17:03:50
Job time : 14.113 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-2  
Perfect score: 489  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482	98.6	222	2 Q9L577	Q9L577 streptococc
2	482	98.6	262	2 Q9L576	Q9L576 streptococc
3	482	98.6	415	2 Q9LAY7	Q9LAY7 streptococc
4	475	97.1	416	2 Q9LAY8	Q9LAY8 streptococc
5	473	96.7	225	2 Q9L591	Q9L591 streptococc
6	441	90.2	406	2 Q9LAZ0	Q9LAZ0 streptococc
7	440	90.0	340	2 Q8KQK5	Q8KQK5 streptococc
8	438	89.6	394	2 Q9LAY6	Q9LAY6 streptococc
9	438	89.6	395	2 Q9LAZ1	Q9LAZ1 streptococc
10	427	87.3	194	2 Q9L5B5	Q9L5B5 streptococc
11	427	87.3	218	2 Q6UEB2	Q6UEB2 streptococc
12	427	87.3	233	2 Q9L568	Q9L568 streptococc
13	427	87.3	236	2 Q9L569	Q9L569 streptococc
14	427	87.3	243	2 Q9L564	Q9L564 streptococc
15	427	87.3	243	2 Q9L567	Q9L567 streptococc
16	427	87.3	244	2 Q9L565	Q9L565 streptococc
17	427	87.3	247	2 Q9L566	Q9L566 streptococc
18	427	87.3	249	2 Q9L570	Q9L570 streptococc
19	427	87.3	254	2 Q9L563	Q9L563 streptococc
20	427	87.3	401	2 Q9LAZ2	Q9LAZ2 streptococc
21	423	86.5	246	2 Q9L578	Q9L578 streptococc
22	418	85.5	255	2 Q9L581	Q9L581 streptococc
23	418	85.5	255	2 Q9L5B6	Q9L5B6 streptococc
24	407	83.2	237	2 Q9L592	Q9L592 streptococc
25	407	83.2	395	2 Q9LAY9	Q9LAY9 streptococc
26	405	82.8	207	2 Q8GNS9	Q8GNS9 streptococc
27	404	82.6	393	2 Q9LAZ3	Q9LAZ3 streptococc
28	352.5	72.1	417	2 Q9LAY3	Q9LAY3 streptococc
29	336.5	68.8	619	2 Q54972	Q54972 streptococc
30	336.5	68.8	619	2 Q8DR10	Q8DR10 streptococc
31	335.5	68.6	739	2 Q9RQT4	Q9RQT4 streptococc

32	335.5	68.6	820	2 Q9RQT1	Q9RQT1 streptococc
33	335.5	68.6	929	2 Q9KK19	Q9KK19 streptococc
34	335.5	68.6	929	2 Q9ZAY5	Q9ZAY5 streptococc
35	328.5	67.2	415	2 Q9LAY1	Q9LAY1 streptococc
36	323.5	66.2	99	2 Q8KQK4	Q8KQK4 streptococc
37	319.5	65.3	437	2 Q9LAY4	Q9LAY4 streptococc
38	316.5	64.7	249	2 Q9L575	Q9L575 streptococc
39	313.5	64.1	426	2 Q9L575	Q9L575 streptococc
40	309.5	63.3	224	2 Q8GNS8	Q8GNS8 streptococc
41	304.5	62.3	395	2 Q9LAY2	Q9LAY2 streptococc
42	304.5	62.3	408	2 Q9LAY0	Q9LAY0 streptococc
43	300	61.3	869	2 Q9KK27	Q9KK27 streptococc
44	167	34.2	246	2 Q9L5B4	Q9L5B4 streptococc
45	164	33.5	653	2 Q34097	Q34097 streptococc

ALIGNMENTS

RESULT 1  
Q9L577 PRELIMINARY; PRT; 222 AA.  
AC Q9L577;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones";  
RL J. Clin. Microbiol. 38:3663-3669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=130;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF255550; AAF68103.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 222 222  
SQ SEQUENCE 222 AA; 24558 MW; 6D7EB7842FE9F2A6 CRC64;

Query Match 98.6%; Score 482; DB 2; Length 222;  
Best Local Similarity 99.0%; Pred. No. 1.8e-22;  
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	LKEIDSESDYVKEGLRAPLQSELDLQAKLSKLEELSDKIDELDAEIAKLEKXNVEDFK 60
Db	25	LKEIDSESDYVKEGLRAPLQSELDLQAKLSKLEELSDKIDELDAEIAKLEKXNVEDFK 84
Qy	61	NSNGSQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEPE 100
Db	85	NSNGSQAEQYRAAAEEDLAAKQAELEKTEADLKAVHEPE 124

RESULT 2  
Q9L576 PRELIMINARY; PRT; 262 AA.  
AC Q9L576;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;

```
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RX STRAIN=232;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255551; AAF68104.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1
FT NON TER 262
FT NON TER 262
SQ SEQUENCE 262 AA; 29012 MW; 32C769099466A584 CRC64;

Query Match 98.6%; Score 482; DB 2; Length 262;
Best Local Similarity 99.0%; Pred. No. 2.1e-22;
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 60
|||||
Db 65 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 124
|||||

Qy 61 NSNGEAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||
Db 125 NSNGEAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 164
|||||

RESULT 3
Q9LAY7
ID Q9LAY7 PRELIMINARY; PRT; 415 AA.
AC Q9LAY7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG6692;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071808; AAF27704.1; -.
FT NON TER 415
FT NON TER 415
SQ SEQUENCE 415 AA; 45593 MW; 41375ACBFAL0FA46 CRC64;

Query Match 98.6%; Score 482; DB 2; Length 415;
Best Local Similarity 99.0%; Pred. No. 3.1e-22;
Matches 99; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 60
|||||
Db 229 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 288
|||||

Qy 61 NSNGEAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||

RESULT 4
Q9LAY8
ID Q9LAY8 PRELIMINARY; PRT; 416 AA.
AC Q9LAY8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8838;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071807; AAF27703.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 416
FT NON TER 416
SQ SEQUENCE 416 AA; 45987 MW; 950C8858BC6B12C7 CRC64;

Query Match 97.1%; Score 475; DB 2; Length 416;
Best Local Similarity 98.0%; Pred. No. 8.4e-22;
Matches 98; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 60
|||||
Db 229 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSELSKIDELDAEIAKLEKNVEDFK 288
|||||

Qy 61 NSNGEAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||
Db 289 NSNGEAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 328
|||||

RESULT 5
Q9L591
ID Q9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI99;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI99;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254258; AAF68093.1; -.

```



```
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
PT NON_TER 225
SQ SEQUENCE 225 AA; 24835 MW; F878A7618B72A692 CRC64;

Query Match 96.78; Score 473; DB 2; Length 225;
Best Local Similarity 97.08; Pred. No. 6.4e-22;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
   |||||
Db 34 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 93
   |||||

QY 61 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||
Db 94 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVNEPE 133
   |||||

RESULT 6
Q9LAZO PRELIMINARY; PRT; 406 AA.
ID Q9LAZO STRAIN=DL6A;
AC Q9LAZO MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 90.2%; Score 441; DB 2; Length 406;
Best Local Similarity 91.0%; Pred. No. 1e-19;
Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
   |||||
Db 213 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 272
   |||||

QY 61 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||
Db 273 NSDGEQAGQYLAABEDLIAKAELEQTADLKKAVNEPE 312
   |||||

RESULT 7
Q8KQK5 PRELIMINARY; PRT; 340 AA.
ID Q8KQK5 STRAIN=L81905;
AC Q8KQK5 MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 89.6%; Score 438; DB 2; Length 394;
Best Local Similarity 91.0%; Pred. No. 1.5e-19;
Matches 91; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
   |||||
Db 213 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 272
   |||||

QY 61 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||
Db 273 NSDGEQAGQYLAABEDLIAKAELEKTEADLKKAVDEPE 312
   |||||
```

```
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
RA DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
immunization with DNA vaccines against Streptococcus pneumoniae";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 340
SQ SEQUENCE 340 AA; 38023 MW; EE07ECF00B1FBD57 CRC64;

Query Match 90.0%; Score 440; DB 2; Length 340;
Best Local Similarity 91.0%; Pred. No. 9.9e-20;
Matches 91; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
   |||||
Db 197 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 256
   |||||

QY 61 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||
Db 257 NSDGEQAGQYLAABEDLIAKAELEKTEADLKKAVNEPE 296
   |||||

RESULT 8
Q9LAY6 PRELIMINARY; PRT; 394 AA.
ID Q9LAY6 STRAIN=L81905;
AC Q9LAY6 MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27705.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 394
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C88FAA CRC64;

Query Match 89.6%; Score 438; DB 2; Length 394;
Best Local Similarity 91.0%; Pred. No. 1.5e-19;
Matches 91; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
   |||||
Db 213 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 272
   |||||

QY 61 NSNGEQAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||
Db 273 NSDGEQAGQYLAABEDLIAKAELEKTEADLKKAVDEPE 312
   |||||
```

```
RESULT 9
Q9LAZ1 PRELIMINARY; PRT; 395 AA.
ID Q9LAZ1
AC Q9LAZ1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9739.
RX MEDLINE=20448953; PubMed=10992499;
DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27700.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECAC41DB7F95 CRC64;

Query Match 89.6%; Score 438; DB 2; Length 395;
Best Local Similarity 91.0%; Pred. No. 1.5e-19;
Matches 91; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 273 NSDGEQAGYLAAREEDLIANKAELEKAEADLKKAVDEPE 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
Q9LSB5 PRELIMINARY; PRT; 194 AA.
ID Q9LSB5
AC Q9LSB5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253407; AAF67355.1; -.
```

```
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
SQ SEQUENCE 194 AA; 21116 MW; E68189FCA2B244F8 CRC64;

Query Match 87.3%; Score 427; DB 2; Length 194;
Best Local Similarity 87.0%; Pred. No. 3.8e-19;
Matches 87; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 55 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 115 NSDGEQAEQYLAAREEDLAAKQAELEKTEADLKKAVDEPE 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
Q6UEB2 PRELIMINARY; PRT; 218 AA.
ID Q6UEB2
AC Q6UEB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URSP2;
PubMed=14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;
RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;
RT "Epitope mapping of a protective monoclonal antibody against
Pneumocystis carinii with shared reactivity to Streptococcus
pneumoniae surface antigen PspA.";
RL Infect. Immun. 72:1548-1556(2004).
DR EMBL; AY371665; AAR20918.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 218
SQ SEQUENCE 218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;

Query Match 87.3%; Score 427; DB 2; Length 218;
Best Local Similarity 87.0%; Pred. No. 4.2e-19;
Matches 87; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 LKEIDSESDYVKEGLRAPLQSELDAAKQAKSLKLELSKIDELDAIAKLEKNVEDFK 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 NSDGEQAEQYLAAREEDLAAKQAELEKTEADLKKAVDEPE 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
Q9LS68 PRELIMINARY; PRT; 233 AA.
ID Q9LS68
AC Q9LS68;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
```

RP	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RX	MEDLINE=20472698; PubMed=11015380;	
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;	
RT	"Pneumococcal pspp sequence types of prevalent multiresistant	
RT	pneumococcal strains in the United States and of internationally	
RT	disseminated clones.";	
RL	J. Clin. Microbiol. 38:3663-3669(2000).	
RN	(2)	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RA	Beall B.W.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF255902; AAF70092.1; -	
DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON TER 1	
FT	NON TER 233	
SQ	SEQUENCE 233 AA; 24514 MW; D5C494019C45BFE2 CRC64;	
Query Match 87.3%; Score 427; DB 2; Length 233;		
Best Local Similarity 87.0%; Pred. No. 4.5e-19;		
Matches 87; Conservative 8; Mismatches 5; Indels 0; Gaps 0;		
Qy	1 LKTEIDSDSDYVKEGLRAPLQSELDAAQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 60	
Db	:     :     :     :     :     :     :     :     :	
Qy	28 LKTEIDSDSDYVKEGLRAPLQSELDAAQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 87	
Db	:     :     :     :     :     :     :     :	
Qy	61 NSNGEQAEQYRAAAEEDLAAQAELEKTEADLKKAVHEPE 100	
Db	:     :     :     :     :     :     :	
Qy	88 NSDGEAQEQLVAAKKOLDAAKALENTADLKKAVDEPE 127	
Db	:     :     :     :     :     :     :	
RESULT 13		
Q9L569	PRELIMINARY; PRT; 236 AA.	
ID	Q9L569	
AC	Q9L569;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	Pspp (Fragment).	
GN	Name=pspp;	
OS	Streptococcus pneumoniae.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1313;	
RN	(1)	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=177;	
RX	MEDLINE=20472698; PubMed=11015380;	
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;	
RT	"Pneumococcal pspp sequence types of prevalent multiresistant	
RT	pneumococcal strains in the United States and of internationally	
RT	disseminated clones.";	
RL	J. Clin. Microbiol. 38:3663-3669(2000).	
RN	(2)	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=177;	
RA	Beall B.W.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF255901; AAF70091.1; -	
DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON TER 1	
FT	NON TER 236	
SQ	SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;	
Query Match 87.3%; Score 427; DB 2; Length 236;		
Best Local Similarity 87.0%; Pred. No. 4.5e-19;		
Matches 87; Conservative 8; Mismatches 5; Indels 0; Gaps 0;		
Qy	1 LKTEIDSDSDYVKEGLRAPLQSELDAAQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 60	
Db	:     :     :     :     :     :     :     :	
Qy	49 LKTEIDSDSDYVKEGLRAPLQSELDAAQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 108	
Db	:     :     :     :     :     :     :	

Search completed: June 18, 2005, 17:01:33  
Job time : 60.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-3  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003Bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	472	96.3	550	8	Adk48356 Streptococ
2	472	96.3	550	8	Adr95223 Novel S.
3	429	87.6	183	2	Aaw14570 Streptococ
4	429	87.6	183	7	Abw02604 Bg9739c p
5	429	87.6	8991	6	Abu08487 S. pneumo
6	419	85.5	194	2	Aaw14584 Streptococ
7	419	85.5	194	7	Abw02618 Dbl6ac pn
8	412	84.1	168	7	Abw02609 L81905c p
9	396.5	80.9	167	2	Aaw14575 Streptococ
10	385	78.6	166	2	Aaw14568 Streptococ
11	385	78.6	166	7	Abw02602 Bg8743c p
12	377.5	77.0	185	2	Aaw14566 Streptococ
13	377.5	77.0	185	7	Abw02600 Ac94c pne
14	348.5	71.1	204	2	Aaw14571 Streptococ
15	348.5	71.1	204	7	Abw02605 Bf1019c p
16	332.5	67.9	198	7	Abw02615 Rxl c pneu
17	332.5	67.9	315	2	Aay04375 Streptococ
18	332.5	67.9	619	2	Aar63437 Pneumococ
19	332.5	67.9	619	2	Aar87598 Pneumococ
20	332.5	67.9	619	2	Aar86911 Pneumococ
21	332.5	67.9	619	2	Aay41838 Streptococ
22	332.5	67.9	619	5	Aae18782 S. pneumo
23	332.5	67.9	619	6	Abu45778 Protein e
24	332.5	67.9	619	8	Ado52126 Streptococ
25	332.5	67.9	648	2	Aaw70336 Pneumococ

26	332.5	67.9	648	2	AAW62274	Aaw62274 Streptoco
27	332.5	67.9	648	2	AAW41837	Aay41837 Streptoco
28	332.5	67.9	648	2	AAW87879	Aaw87879 A. pneumoc
29	332.5	67.9	653	2	AAW92456	Aaw92456 S. pneumo
30	332.5	67.9	684	2	AAW73912	Aar73912 Streptoco
31	331.5	67.7	170	7	ABW02614	Abw02614 Rct135c p
32	331.5	67.7	181	7	ABW02596	Abw02596 0922134c
33	331.5	67.7	198	2	AAW14581	Aaw14581 Streptoco
34	331.5	67.7	865	6	ABU08489	Abu08489 S. pneumo
35	331.5	67.7	929	2	AAW14593	Aaw14593 Streptoco
36	331.5	67.7	929	2	AAW43384	Aay43384 S. pneumo
37	328.5	67.0	188	2	AAW14580	Aaw14580 Streptoco
38	328.5	67.0	188	7	ABW02613	Abw02613 Rct139c p
39	320.5	65.4	195	2	AAW14591	Aaw14591 Streptoco
40	320.5	65.4	195	7	ABW02625	Abw02625 Wu2c pneu
41	319.5	65.2	588	6	ABU08491	Abu08491 Coiled co
42	319.5	65.2	589	2	AAW43392	Aay43392 PspC alph
43	317.5	64.8	204	2	AAW14578	Aaw14578 Streptoco
44	317.5	64.8	204	7	ABW02612	Abw02612 Rct133c p
45	317	64.7	180	2	AAW14562	Aaw14562 Streptoco

## ALIGNMENTS

RESULT 1  
ADK48356  
ID ADK48356 standard; protein; 550 AA.  
XX AC ADK48356;  
XX DT 20-MAY-2004 (first entry)  
XX DE Streptococcus pneumoniae protein, Seq ID No 4871.  
XX KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
XX OS Streptococcus pneumoniae.  
XX PN US6699703-B1.  
XX PD 02-MAR-2004.  
XX PF 26-MAY-2000; 2000US-00583110.  
XX PR 02-JUL-1997; 97US-0051553P.  
XX PR 12-MAY-1998; 98US-0085131P.  
XX PR 30-JUN-1998; 98US-00107433.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;  
XX WPI; 2004-212399/20.  
XX N-PSDB; ADK45695.  
XX New nucleic acid molecules and polypeptides useful for diagnosing,  
XX preventing and treating pathological conditions resulting from bacterial  
XX infection, e.g. Streptococcus pneumoniae infection, and in drug  
XX screening.  
XX Disclosure; SEQ ID NO 4871; 301pp; English.  
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
XX and polypeptides. The nucleic acids and proteins are useful for  
XX diagnosing, preventing and treating pathological conditions resulting  
XX from bacterial infection, such as S. pneumoniae infection. These may also  
XX be used for drug screening procedures. The present sequence represents a  
XX Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
XX data for this patent did not appear in the printed specification but was  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html.

```
SQ Sequence 550 AA;
Query Match 96.3%; Score 472; DB 8; Length 550;
Best Local Similarity 97.0%; Pred. No. 4.4e-35;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 1 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAIAKLEKNVEDFK 60
Db 144 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAIAKLEKNVEDFK 203
Oy 61 NSNGEAEQYRAAAGEDIAAKQAELEKTEADLKKAHVHEPE 100
Db 204 NSNGEAEQYRAAABEDIAAKQAELEKTEADLKKAHVNEPE 243

RESULT 2
ADR95223
ID ADR95223 standard; protein; 550 AA.
XX
AC ADR95223;
XX
DT 16-DEC-2004 (first entry)
XX
DE Novel S. pneumoniae protein sequence, SEQ ID 3858.
XX
KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
KW bacterial infection.
XX
OS Streptococcus pneumoniae.
XX
PN US6800744-B1.
XX
PD 05-OCT-2004.
XX
PF 30-JUN-1998; 98US-00107433.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
WPI; 2004-697205/68.
DR N-PSDB; ADR92620.
XX
New isolated nucleic acid encoding a Streptococcus pneumoniae
PT polypeptide, useful for diagnosing, preventing and/or treating
PT pathological conditions resulting from the bacterial infection.
XX
PS Disclosure; SEQ ID NO 3858; 151pp; English.
XX
The invention relates to an isolated nucleic acid comprising a sequence
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
CC hybridisable under high stringency conditions to the nucleotide sequence.
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
CC Also included are a recombinant expression vector comprising the isolated
CC nucleic acid cited above operably linked to a transcription regulatory
CC element, a cell comprising the recombinant expression vector and a probe
CC comprising at least 20 consecutive nucleotides of the nucleotide
CC sequences as cited above. The methods and compositions of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC pathological conditions resulting from bacterial infection by
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
CC otitis media. The present sequence is one of the 2603 disclosed S.
CC pneumoniae protein sequences. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
```

```
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.
XX
SQ Sequence 550 AA;
Query Match 96.3%; Score 472; DB 8; Length 550;
Best Local Similarity 97.0%; Pred. No. 4.4e-35;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 1 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAIAKLEKNVEDFK 60
Db 144 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAIAKLEKNVEDFK 203
Oy 61 NSNGEAEQYRAAAGEDIAAKQAELEKTEADLKKAHVHEPE 100
Db 204 NSNGEAEQYRAAABEDIAAKQAELEKTEADLKKAHVNEPE 243

RESULT 3
AAW14570
ID AAW14570 standard; protein; 183 AA.
XX
AC AAW14570;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Bg9739.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
WPI; 1997-202002/18.
XX
Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
Example 6; Fig 13; 296pp; English.
XX
This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 183 AA;
Query Match 87.6%; Score 429; DB 2; Length 183;
Best Local Similarity 88.0%; Pred. No. 1.1e-31;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
QY 61 NSNGEAEQYRAAAGEDLIAKAELEKAEADLKKAVHEPE 100
DB 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDEPE 100

RESULT 4
ID ABW02604 standard; protein; 183 AA.
XX
AC ABW02604;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bg9739c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
DR WPI; 2003-862841/80.
XX
PT Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 50; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies), or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bg9739c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 183 AA;
Query Match 87.6%; Score 429; DB 7; Length 183;
Best Local Similarity 88.0%; Pred. No. 1.1e-31;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
QY 61 NSNGEAEQYRAAAGEDLIAKAELEKAEADLKKAVHEPE 100
DB 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDEPE 100
```

```
DB 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDEPE 100
RESULT 5
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
AC ABU08487;
XX
DT 24-JUN-2003 (first entry)
XX
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
OS Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..8991
FT /note="All Xaa residues within this sequence are
FT unknown"
XX
PN US6500613-B1.
XX
PD 31-DEC-2002.
XX
PF 16-SEP-1996; 96US-00714741.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 2003-361534/34.
XX
PT Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
PS Disclosure; Col 145-188; 186pp; English.
XX
CC The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents S. pneumoniae
CC PspA protein
XX
SQ Sequence 8991 AA;
Query Match 87.6%; Score 429; DB 6; Length 8991;
Best Local Similarity 88.0%; Pred. No. 1.1e-29;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
DB 5139 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 5138
QY 61 NSNGEAEQYRAAAGEDLIAKAELEKAEADLKKAVHEPE 100
DB 5199 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDEPE 5238
RESULT 6
AAW14584
```

```
ID AAW14584 standard; protein; 194 AA.
XX AC AAW14584;
XX
XX 17-OCT-2003 (revised)
DT DT 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
DE
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Db16.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 61
FT /note= "unidentified amino acid"
FT
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Db16.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 194 AA;
SQ
Query Match 85.5%; Score 419; DB 2; Length 194;
Best Local Similarity 87.9%; Pred. No. 9.9e-31;
Matches 87; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAQAKLSKLELSKIDELDAIAKLEKNVEDFK 60
DB 1 LKEIDESDSEYVKEGFRAPLQSELDAAQAKLSKLELSKIDELDAIAKLEKNVEDFK 60
QY 61 NSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAHVHP 99
DB 61 XSDGEQAGQYLAAREEDLIAKAELEQTEADLKKAHVHP 99
RESULT 8
ABW02609
ID ABW02609 standard; protein; 168 AA.
XX AC ABW02609;
XX
XX 12-FEB-2004 (first entry)
DT
XX
```

```
DT 12-FEB-2004 (first entry)
XX
XX Db16ac pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 1. .194
FT /note= "Xaa = Unknown amino acid"
FT
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 64; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
XX Sequence 194 AA;
SQ
Query Match 85.5%; Score 419; DB 7; Length 194;
Best Local Similarity 87.9%; Pred. No. 9.9e-31;
Matches 87; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAQAKLSKLELSKIDELDAIAKLEKNVEDFK 60
DB 1 LKEIDESDSEYVKEGFRAPLQSELDAAQAKLSKLELSKIDELDAIAKLEKNVEDFK 60
QY 61 NSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAHVHP 99
DB 61 XSDGEQAGQYLAAREEDLIAKAELEQTEADLKKAHVHP 99
RESULT 8
ABW02609
ID ABW02609 standard; protein; 168 AA.
XX AC ABW02609;
XX
XX 12-FEB-2004 (first entry)
DT
```



DE L81905c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1.168  
 FT /note= "Xaa = Unknown amino acid"  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 55; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is L81905c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 168 AA;  
 Query Match 84.1%; Score 412; DB 7; Length 168;  
 Best Local Similarity 86.0%; Pred. No. 3.7e-30;  
 Matches 86; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAETIAKLEKNVEDPK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAETIAKLEKNVEDPK 60  
 QY 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100  
 DB 61 NSDGEAGQYLAAREEDLIAKXAEADLKAVDEPE 100  
 RESULT 9  
 AAW14575  
 ID AAW14575 standard; protein; 167 AA.  
 XX  
 AC AAW14575;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain L81905.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 37  
 FT /note= "unidentified amino acid"  
 FT Misc-difference 41  
 FT /note= "unidentified amino acid"  
 FT Misc-difference 83  
 FT /note= "unidentified amino acid"  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 DR WPI; 1997-202002/18.  
 XX  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 PS Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain AAL81905.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 167 AA;  
 Query Match 80.9%; Score 396.5; DB 2; Length 167;  
 Best Local Similarity 85.0%; Pred. No. 9.8e-29;  
 Matches 85; Conservative 4; Mismatches 10; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAETIAKLEKNVEDPK 60  
 DB 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAETIAKLEKNVEDPK 59  
 QY 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100  
 DB 60 NSDGEAGQYLAAREEDLIAKXAEADLKAVDEPE 99  
 RESULT 10  
 AAW14568  
 ID AAW14568 standard; protein; 166 AA.  
 XX  
 AC AAW14568;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.

```
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Bg8743.
OS
XX WO9709994-A1.
FN
XX
XX 20-MAR-1997.
PD
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
PR
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
DR
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
PT
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 166 AA;
SQ
Query Match 78.6%; Score 385; DB 2; Length 166;
Best Local Similarity 80.0%; Pred. No. 1.1e-27;
Matches 80; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLEKVDGDFP 60
QY 61 NSNGEEAQYRAAGEDLAAKQAELEKTEADLKKAVHEPE 100
DB 61 NSDGEQAGQYLVAEAKDLDAKEAEELGNTGADLKKAVDEPE 100
RESULT 11
ABW02602
ID ABW02602 standard; protein; 166 AA.
XX
XX ABW02602;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX Bg8743c pneumococcal surface protein A (PspA) central region.
DE
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
OS
XX US6592876-B1.
FN
XX
XX 15-JUL-2003.
PD
```

```
XX 15-SEP-1995; 95US-00529055.
PF
XX 20-APR-1993; 93US-00048896.
PR
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
PA
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 48; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspAs) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antigenic,
XX immunological or vaccine compositions, for eliciting antibodies, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Bg8743c
XX pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
XX Sequence 166 AA;
SQ
Query Match 78.6%; Score 385; DB 7; Length 166;
Best Local Similarity 80.0%; Pred. No. 1.1e-27;
Matches 80; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKNVEDFK 60
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLEKVDGDFP 60
QY 61 NSNGEEAQYRAAGEDLAAKQAELEKTEADLKKAVHEPE 100
DB 61 NSDGEQAGQYLVAEAKDLDAKEAEELGNTGADLKKAVDEPE 100
RESULT 12
AAW14566
ID AAW14566 standard; protein; 185 AA.
XX
XX AAW14566;
AC
XX
XX 17-OCT-2003 (revised)
DT
XX 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
DE
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Ac94.
OS
XX WO9709994-A1.
FN
XX
XX 20-MAR-1997.
PD
XX
XX 16-SEP-1996; 96WO-US014819.
PF
XX
XX 15-SEP-1995; 95US-00529055.
PR
XX
```

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 77.0%; Score 377.5; DB 2; Length 185;  
 Best Local Similarity 80.2%; Pred. No. 6.3e-27;  
 Matches 81; Conservative 7; Mismatches 12; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYVKEGLRAPLQSELDVKAQKLSKLELSKDIDELDAETAK-LKKNVDF 59  
 DB 1 LKEIDESDSYVKEGLRVPLQSELDVKAQKLSKLELSKDIDELDAETAKLKQVDF 60  
 QY 60 KNSNGEAEQYRAAGEDLAAQAELEKTEADLKAVHEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKAVNEPE 101  
 RESULT 13  
 AEW02600  
 ID AEW02600 standard; protein; 185 AA.  
 AC AEW02600;  
 XX 12-FEB-2004 (first entry)  
 XX Ac94c pneumococcal surface protein A (PspA) central region.  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 OS US6592876-B1.  
 XX US6592876-B1.  
 XX 15-JUL-2003.  
 XX 15-SEP-1995; 95US-00529055.  
 XX 20-APR-1993; 93US-00048896.  
 XX 06-JUN-1995; 95US-00465746.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,

PT comprises at least two different full length isolated gene encoding  
 XX pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 46; 121pp; English.  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 77.0%; Score 377.5; DB 7; Length 185;  
 Best Local Similarity 80.2%; Pred. No. 6.3e-27;  
 Matches 81; Conservative 7; Mismatches 12; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYVKEGLRAPLQSELDVKAQKLSKLELSKDIDELDAETAK-LKKNVDF 59  
 DB 1 LKEIDESDSYVKEGLRVPLQSELDVKAQKLSKLELSKDIDELDAETAKLKQVDF 60  
 QY 60 KNSNGEAEQYRAAGEDLAAQAELEKTEADLKAVHEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKAVNEPE 101  
 RESULT 14  
 AAW14571  
 ID AAW14571 standard; protein; 204 AA.  
 XX AAW14571;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Ef1019.  
 OS WO9709994-A1.  
 XX WO9709994-A1.  
 XX 20-MAR-1997.  
 XX 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX  
SQ Sequence 204 AA;  
Query Match 71.1%; Score 348.5; DB 2; Length 204;  
Best Local Similarity 74.0%; Pred. No. 3.3e-24;  
Matches 74; Conservative 11; Mismatches 14; Indels 1; Gaps 1;  
Qy 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60  
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEQDLKAAE 60  
Qy 61 NSNGEEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100  
Db 61 ENNVVE-DYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

RESULT 15  
ABW02605  
ID ABW02605 standard; protein; 204 AA.  
XX  
AC ABW02605;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Ef1019c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
XX WPI; 2003-862841/80.

XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 51; 121pp; English.

XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Ef1019c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 204 AA;

Query Match 71.1%; Score 348.5; DB 7; Length 204;  
Best Local Similarity 74.0%; Pred. No. 3.3e-24;  
Matches 74; Conservative 11; Mismatches 14; Indels 1; Gaps 1;  
Qy 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60  
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEQDLKAAE 60  
Qy 61 NSNGEEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100  
Db 61 ENNVVE-DYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

Search completed: June 18, 2005, 16:51:19  
Job time : 73.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-3  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKKAHVEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	490	100.0	100	4	US-09-147-875A-3
2	479	97.8	100	4	US-09-147-875A-2
3	472	96.3	550	4	US-09-583-110-4871
4	472	96.3	550	4	US-09-107-433-3858
5	469.5	95.8	101	2	US-08-710-749-2
6	461.5	94.2	101	2	US-08-710-749-1
7	435	88.8	100	4	US-09-147-875A-5
8	433	88.4	100	4	US-09-147-875A-4
9	429	87.6	183	4	US-08-529-055-50
10	429	87.6	8991	4	US-08-714-741-32
11	428	87.3	98	4	US-09-147-875A-1
12	422.5	86.2	101	2	US-08-710-749-3
13	419	85.5	194	4	US-08-529-055-64
14	417.5	85.2	101	2	US-08-710-749-4
15	412	84.1	168	4	US-08-529-055-55
16	410.5	83.8	99	2	US-08-710-749-9
17	409	83.5	100	4	US-09-147-875A-6
18	398.5	81.3	101	2	US-08-710-749-5
19	396	80.8	100	4	US-09-147-875A-8
20	385	78.6	166	4	US-08-529-055-48
21	384.5	78.5	101	4	US-09-147-875A-9
22	383.5	78.3	101	2	US-08-710-749-7
23	377.5	77.0	185	4	US-08-529-055-46
24	370	75.5	100	4	US-09-147-875A-7
25	367	74.9	102	2	US-08-710-749-8
26	357.5	73.0	101	2	US-08-710-749-6
27	348.5	71.1	99	2	US-08-710-749-10

28	348.5	71.1	99	4	US-09-147-875A-11	Sequence 11, Appl
29	348.5	71.1	204	4	US-08-529-055-51	Sequence 51, Appl
30	339	69.2	100	4	US-09-147-875A-12	Sequence 12, Appl
31	332.5	67.9	99	2	US-08-710-749-11	Sequence 11, Appl
32	332.5	67.9	198	4	US-08-529-055-61	Sequence 61, Appl
33	332.5	67.9	619	1	US-08-465-746-2	Sequence 2, Appl
34	332.5	67.9	619	1	US-08-214-164-2	Sequence 2, Appl
35	332.5	67.9	619	2	US-08-467-852A-3	Sequence 3, Appl
36	332.5	67.9	619	2	US-08-246-636-2	Sequence 2, Appl
37	332.5	67.9	619	2	US-08-247-491A-3	Sequence 3, Appl
38	332.5	67.9	619	2	US-08-319-795-2	Sequence 2, Appl
39	332.5	67.9	619	2	US-08-468-985-2	Sequence 2, Appl
40	332.5	67.9	619	3	US-08-312-949-2	Sequence 2, Appl
41	332.5	67.9	648	1	US-08-072-070-2	Sequence 2, Appl
42	332.5	67.9	648	1	US-08-469-434-2	Sequence 2, Appl
43	332.5	67.9	648	1	US-08-214-222-2	Sequence 2, Appl
44	332.5	67.9	648	2	US-08-467-852A-2	Sequence 2, Appl
45	332.5	67.9	648	2	US-08-468-718-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-147-875A-3  
; Sequence 3, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-3

Query Match 100.0%; Score 490; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.6e-39;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	60
DB	1	LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	60

QY	61	NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVEPE	100
DB	61	NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVEPE	100

RESULT 2  
US-09-147-875A-2  
; Sequence 2, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-2

Query Match 97.8%; Score 479; DB 4; Length 100;

Best Local Similarity	98.0%;	Pred. No. 1.7e-38;			
Matches	98;	Conservative	1;	Mismatches	1;
				Indels	0;
				Gaps	0;
QY	1	LKEIDSDSDYVKEGIRAPLQSELDAAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	60		
Db	1	LKEIDSDSDYVKEGIRAPLQSELDAAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	60		
QY	61	NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKAVHEPE	100		
Db	61	NSNGEAEQYRAAABEDLAAKQAELEKTEADLKAVHEPE	100		
RESULT 3					
US-09-583-110-4871					
; Sequence 4871, Application US/09583110					
; Patent No. 6699703					
; GENERAL INFORMATION:					
; APPLICANT: Lynn Doucette-Stamm et al.					
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus					
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics					
; FILE REFERENCE: PATH00-07A					
; CURRENT APPLICATION NUMBER: US/09/583,110					
; CURRENT FILING DATE: 2000-05-26					
; PRIOR APPLICATION NUMBER: US 09/107,433					
; PRIOR FILING DATE: 1998-06-30					
; PRIOR APPLICATION NUMBER: US 60/085,131					
; PRIOR FILING DATE: 1998-05-12					
; PRIOR APPLICATION NUMBER: US 60/051,553					
; PRIOR FILING DATE: 1997-07-02					
; NUMBER OF SEQ ID NOS: 5322					
; SEQ ID NO 4871					
; LENGTH: 550					
; TYPE: PRT					
; ORGANISM: Streptococcus pneumoniae					
US-09-583-110-4871					
Query Match 96.3%; Score 472; DB 4; Length 550;					
Best Local Similarity 97.0%; Pred. No. 6.5e-37;					
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;					
QY	1	LKEIDSDSDYVKEGIRAPLQSELDAAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	60		
Db	144	LKEIDSDSDYVKEGIRAPLQSELDAAKQAKLSKEELSDKIDELDAETAKLEKNVEDFK	203		
QY	61	NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKAVHEPE	100		
Db	204	NSNGEAEQYRAAABEDLAAKQAELEKTEADLKAVHEPE	243		
RESULT 4					
US-09-107-433-3858					
; Sequence 3858, Application US/09107433					
; Patent No. 6800744					
; GENERAL INFORMATION:					
; APPLICANT: Lynn A Doucette-Stamm and David Bush					
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID					
; TITLE OF INVENTION: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE					
; TITLE OF INVENTION: THERAPEUTICS					
; NUMBER OF SEQUENCES: 5206					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION					
; STREET: 100 Beaver Street					
; CITY: Waltham					
; STATE: Massachusetts					
; COUNTRY: USA					
; ZIP: 02354					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: CD-ROM ISO9660					
; COMPUTER: <Unknown>					
; OPERATING SYSTEM: <Unknown>					
; SOFTWARE: <Unknown>					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/09/107,433					

```

; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-2

Query Match          95.8%; Score 469.5; DB 2; Length 101;
Best Local Similarity 97.0%; Pred. No. 1.4e-37;
Matches 98; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRAPLQSEL-DAKQAKLSKLELSDKIDELDAEIAKLEKNVEDF 59
Db 1 LKEIDSDSDYVKEGLRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEKNVEDF 60

Qy 60 KNSNGERAEQVRAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 KNSNGEAEQVRAAGEDLAAKQAELEKTEADLKKAVNEPE 101

RESULT 6
US-08-710-749-1
; Sequence 1, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/710,749
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-1

Query Match          94.2%; Score 461.5; DB 2; Length 101;
Best Local Similarity 96.0%; Pred. No. 8.1e-37;
Matches 97; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRAPLQSEL-DAKQAKLSKLELSDKIDELDAEIAKLEKNVEDF 59

```

```

Db 1 LKEIDSDSDYVYKEGLRAPLQSELDLDAQAQKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Qy 60 KNSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAVHEPE 100
Db 61 KNSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAVNEPE 101

RESULT 7
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match 88.4%; Score 433; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 2.6e-34;
Matches 90; Conservative 5; Mismatches 5; Indels 0; Gaps 0

Qy 1 LKEIDSDSDYVYKEGLRAPLQSELDLDAQAQKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSDSDYVYKEGERAPLQSELDLDAQAQKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Qy 61 NNSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100

RESULT 8
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

Query Match 88.4%; Score 433; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 4.1e-34;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0

Qy 1 LKEIDSDSDYVYKEGLRAPLQSELDLDAQAQKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSDSDYVYKEGERAPLQSELDLDAQAQKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Qy 61 NNSNGEAEQYRAAAGEDLAAQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEQAGQYLAABEDLIAKAELEKAEADLKKAVDEPE 100

RESULT 9
US-08-529-055-50
; Sequence 50, Application US/08529055
; Patent No. 6592876

```

```

Db      1 LKEIDSDSEDYVYKGLRAPLQSELDADKAQAKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Oy      60 KNSNGEAEOYRAAAGEDLAQAQAELEKTEADLKAVHEPE 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 KNSNGEAEOYRAAAGEDLAQAQAELEKTEADLKAVNEPE 101
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 7
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match          88.4%; Score 433; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 2.6e-34;
Matches 90; Conservative 5; Mismatches 5; Indels 0; Gaps 0

Oy      1 LKEIDSDSEDYVYKGLRAPLQSELDADKAQAKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Db      1 LKEIDSDSEDYVYKGERAPLQSELDAQAQAKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Oy      61 NNSGGEAEQYRAAAGEDIQAQAQAELEKTEADLKAVHEPE 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 NSDGEAQGYLAABEDLIQAQAQAELEQTEADLKAVHEPE 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

Query Match          88.4%; Score 433; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 4.1e-34;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0

Oy      1 LKEIDSDSEDYVYKGLRAPLQSELDADKAQAKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Db      1 LKEIDSDSEDYVYKGERAPLQSELDADKAQAKLSLELSDKIDELDAEIAKLEKNVEDFK 60
Oy      61 NNSGGEAEQYRAAAGEDIQAQAQAELEKTEADLKAVHEPE 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 NSDGEAQGYLAABEDLIQAQAQAELEKTEADLKAVDEPE 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
US-08-529-055-50
; Sequence 50, Application US/08529055
; Patent No. 6592876

```

```

; GENERAL INFORMATION:
; APPLICANT: Briles, David B.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-50

Query Match 87.6%; Score 429; DB 4; Length 183;
Best Local Similarity 88.0%; Pred. No. 2e-33;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEQAGQYLAAGGEDLIAKKAELEKAEADLKKAVDEPE 100

RESULT 10
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match 87.6%; Score 429; DB 4; Length 8991;
Best Local Similarity 88.0%; Pred. No. 2.4e-31;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 5139 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 5198

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 5199 NSDGEQAGQYLAAGGEDLIAKKAELEKAEADLKKAVDEPE 5238

RESULT 11
US-09-147-875A-1
; Sequence 1, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-1

Query Match 87.3%; Score 428; DB 4; Length 98;
Best Local Similarity 92.0%; Pred. No. 1.2e-33;
Matches 92; Conservative 4; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
```



Db 61 NSDGEQA-QYLAAAEEDL-AKKAELKTEADLKAVHEPE 98

RESULT 12  
US-08-710-749-3  
; Sequence 3, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan.  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-3

Query Match 86.2%; Score 422.5; DB 2; Length 101;  
Best Local Similarity 89.1%; Pred. No. 4.1e-33;  
Matches 90; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 LKEIDESDSEYVKEGRAPLQSEL-DAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDF 59  
|||||  
Db 1 LKEIDESDSEYVKEGRAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKVEDF 60  
|||||

QY 60 KNSGGEAQYRAAAGEDLAAKQAELEKTEADLKAVHEPE 100  
|||||  
Db 61 KNSDGEQAGYLAAGEDIKAAELEDIAKAELEKTEADLKAVDEPE 101  
|||||

RESULT 13  
US-08-529-055-64  
; Sequence 64, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 194 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-64

Query Match 85.5%; Score 419; DB 4; Length 194;  
Best Local Similarity 87.9%; Pred. No. 2e-32;  
Matches 87; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYVKEGRAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60  
|||||  
Db 1 LKEIDESDSEYVKEGRAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKVEDFK 60  
|||||

QY 61 NNGGEAQYRAAAGEDLAAKQAELEKTEADLKAVHEP 99  
|||  
Db 61 XSDGEQAGYLAABEDLIAKAELEKTEADLKAVNEP 99  
|||

RESULT 14  
US-08-710-749-4  
; Sequence 4, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-4

Query Match      85.2%; Score 417.5; DB 2; Length 101;
Best Local Similarity 88.1%; Pred. No. 1.2e-32;
Matches 89; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 LKEIDESDSEYVKEGLRAPLQSEL-DAKQAKLSKLELSKIDELDAEIAKLEKNVEDF 59
   |||||
Db 1 LKEIDESDSEYVKEGERAPLQSELDLDAKQAKLSKLELSKIDELDAEIAKLEKVDVF 60
   |||||

QY 60 KNSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100
   |||||
Db 61 KNSDGEAQGYLAARAEEDLIAKKALEKAEADLKAVDEPE 101
   |||||

RESULT 15
US-08-529-055-55
; Sequence 55, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Iother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: peptide
US-08-529-055-55

Query Match      84.1%; Score 412; DB 4; Length 168;
Best Local Similarity 86.0%; Pred. No. 7.6e-32;
Matches 86; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYVKEGLRAPLQSELDLDAKQAKLSKLELSKIDELDAEIAKLEKNVEDFK 60
   |||||
Db 1 LKEIDESDSEYVKEGFAPLQSELDLDAKQAKLSKLEXSXXXDELDAEIAKLEKXVEDFK 60
   |||||

QY 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100
   |||||
Db 61 NSDGEAQGYLAARAEEDLIAKKALEKAEADLKAVDEPE 100
   |||||
```

Search completed: June 18, 2005, 17:07:05  
Job time : 19.9189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-3  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	490	100.0	100	15	US-10-674-755-3
2	479	97.8	100	15	US-10-674-755-2
3	435	88.8	100	15	US-10-674-755-5
4	433	88.4	100	15	US-10-674-755-4
5	429	87.6	183	15	US-10-299-636-65
6	428	87.3	98	15	US-10-674-755-1
7	419	85.5	194	15	US-10-299-636-79
8	412	84.1	168	15	US-10-299-636-70
9	409	83.5	100	15	US-10-674-755-6
10	396	80.8	100	15	US-10-674-755-8
11	385	78.6	166	15	US-10-299-636-63
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 5, Appli
					Sequence 4, Appli
					Sequence 65, Appli
					Sequence 1, Appli
					Sequence 79, Appli
					Sequence 6, Appli
					Sequence 8, Appli
					Sequence 63, Appli

ALIGNMENTS

RESULT 1  
US-10-674-755-3  
; Sequence 3, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-3

Query Match 100.0%; Score 490; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.4e-33;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDSESDYVKEGLRAPQLSELDKQAKLSKLEELSDKTDLDLDAETAKLEKNVEDPK 60  
Db 1 LKEIDSESDYVKEGLRAPQLSELDKQAKLSKLEELSDKTDLDLDAETAKLEKNVEDPK 60  
QY 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100  
Db 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKAVHEPE 100  
RESULT 2

```
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match      97.8%; Score 479; DB 15; Length 100;
Best Local Similarity 98.0%; Pred. No. 3.6e-32;
Matches 98; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 3
US-10-674-755-3
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      88.8%; Score 435; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 1.5e-28;
Matches 90; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 4
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
```

```
US-10-674-755-4
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match      88.4%; Score 433; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 2.2e-28;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 5
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match      87.6%; Score 429; DB 15; Length 183;
Best Local Similarity 88.0%; Pred. No. 9.3e-28;
Matches 88; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDESDSDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 6
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
```

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-1

Query Match 87.3%; Score 428; DB 15; Length 98;  
Best Local Similarity 92.0%; Pred. No. 5.5e-28;  
Matches 92; Conservative 4; Mismatches 2; Indels 2; Gaps 2;  
QY 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
Db 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
QY 61 NSNGEAEQYRAAAGDLAAKQAELEKTEADLKKAVHEP 100  
Db 61 NSDGEQAGYLAAAEEDLAKKAELEKTEADLKKAVHEP 98

RESULT 7  
US-10-299-636-79  
; Sequence 79, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 79  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (61)  
; OTHER INFORMATION: Xaa at position 61 is unknown  
US-10-299-636-79

Query Match 85.5%; Score 419; DB 15; Length 194;  
Best Local Similarity 87.9%; Pred. No. 6.6e-27;  
Matches 87; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
Db 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
QY 61 NSNGEAEQYRAAAGDLAAKQAELEKTEADLKKAVHEP 99  
Db 61 XSDGEQAGYLAAAEEDLAKKAELEKTEADLKKAVNEP 99

RESULT 8  
US-10-299-636-70  
; Sequence 70, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa at position 38 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (42)  
; OTHER INFORMATION: Xaa at position 42 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (84)  
; OTHER INFORMATION: Xaa at position 84 is unknown  
US-10-299-636-70

Query Match 84.1%; Score 412; DB 15; Length 168;  
Best Local Similarity 86.0%; Pred. No. 2.1e-26;  
Matches 86; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
Db 1 LKEIDESDSEYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60  
QY 61 NSNGEAEQYRAAAGDLAAKQAELEKTEADLKKAVHEP 100  
Db 61 NSDGEQAGYLAAAEEDLAKKAELEKTEADLKKAVDEPE 100

RESULT 9  
US-10-674-755-6  
; Sequence 6, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:

```
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match      83.5%; Score 409; DB 15; Length 100;
Best Local Similarity 86.0%; Pred. No. 2.1e-26;
Matches 86; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEAQYLAABEDLIAXKAEADLKKAVDEPE 100

RESULT 10
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-8

Query Match      80.8%; Score 396; DB 15; Length 100;
Best Local Similarity 83.0%; Pred. No. 2.4e-25;
Matches 83; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIPKLEKNVEYFK 60
Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 LTDAEQTYLAABEKDLADKAELEKTEADLKKAVHEPE 100

RESULT 11
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      78.6%; Score 385; DB 15; Length 166;
Best Local Similarity 80.0%; Pred. No. 3.5e-24;
Matches 80; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 NSDGEAQYLVAAEKDLDAKAEELGNTGADLKKAVDEPE 100

RESULT 12
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      78.5%; Score 384.5; DB 15; Length 101;
Best Local Similarity 81.2%; Pred. No. 2.2e-24;
Matches 82; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 59
Db 1 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKLEKNVEDFK 60
Qy 60 KNSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101

RESULT 13
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
```

; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-61

Query Match 77.0%; Score 377.5; DB 15; Length 185;  
Best Local Similarity 80.2%; Pred. No. 1.6e-23;  
Matches 81; Conservative 7; Mismatches 12; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAETAKLEKNVEDF 59  
Db |||||  
1 LKEIDSDSDYVKEGLRVPLQSELDVQAKLKLKLELSKIDELDAETAKLKDVEDF 60  
QY 60 KNSGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100  
Db |||||  
61 QNSGGYSAULYLAEEKDLVAKKAELEKTEADLKKAVNEPE 101

## RESULT 14

US-10-674-755-7  
; Sequence 7, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-7

Query Match 75.5%; Score 370; DB 15; Length 100;  
Best Local Similarity 78.0%; Pred. No. 3.4e-23;  
Matches 78; Conservative 8; Mismatches 14; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAETAKLEKNVEDF 60  
Db |||||  
1 LKEIDSDSDYVKEGLRAPLQSKLDKAKKLSKLELSKIDEXDKXDELDAETAKLEKDVGDFF 60  
QY 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100  
Db |||||  
61 NSDGEQAGYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100

## RESULT 15

US-10-674-755-11  
; Sequence 11, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-11

Query Match 71.1%; Score 348.5; DB 15; Length 99;  
Best Local Similarity 74.0%; Pred. No. 2e-21;  
Matches 74; Conservative 11; Mismatches 14; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRAPLQSELDKQAKLSKLELSKIDELDAETAKLEKNVEDF 60  
Db |||||  
1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAETAKLEQDKAAE 60  
QY 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100  
Db |||||  
61 ENNVVE-DYFKEGLEKTIKAELEKTEADLKKAVNEPE 99

Search completed: June 18, 2005, 18:00:21  
Job time : 63.963 secs

**This Page Blank (uspto)**



GenCore version S.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-3  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	67.9	619	2 A97887	surface protein ps
2	332.5	67.9	619	2 A41971	surface protein ps
3	134	27.3	744	2 P95013	pneumococcal surfa
4	109	22.2	1269	2 F84730	probable myosin he
5	107.5	21.9	1169	2 A64505	Pil15 homolog - Met
6	106.5	21.7	562	2 G70002	hypothetical prote
7	105	21.4	522	2 G02533	occludin - human
8	104.5	21.3	1957	2 T38077	hypothetical coile
9	103.5	21.1	896	2 S43074	epidermal growth f
10	102	20.8	1138	2 T24635	hypothetical prote
11	102	20.8	1959	1 A33977	myosin heavy chain
12	102	20.8	3488	2 T34418	hypothetical prote
13	100	20.4	1006	2 C70445	ATPase subunit of
14	100	20.4	1319	2 A28313	glued protein - fr
15	99.5	20.3	886	2 H63378	conserved hypothet
16	99.5	20.3	924	2 S06117	myosin heavy chain
17	99.5	20.3	1053	2 A41642	dynactin - chicken
18	99.5	20.3	1156	2 B70356	chromosome assembl
19	99.5	20.3	2007	1 B43402	myosin heavy chain
20	99	20.2	161	2 T48396	tropomyosin TPM2 -
21	99	20.2	1164	2 T24806	hypothetical prote
22	99	20.2	1190	2 E84193	chromosome segrega
23	99	20.2	1361	1 A61231	myosin heavy chain
24	98.5	20.1	157	2 A97703	ATP synthase B cha
25	98	20.0	1938	2 A59293	skeletal myosin he
26	97.5	19.9	281	2 F75216	hypothetical prote
27	97.5	19.9	1976	2 A59252	myosin heavy chain
28	97	19.8	2288	2 T29998	hypothetical prote
29	96.5	19.7	387	2 S57834	fcrA protein precu

ALIGNMENTS

RESULT 1

A97887

surface protein psPA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887

A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:A5007317; PIDN:AAK98925.1; PID: g  
C:Genetics:  
A:Gene: psPA

Query Match 67.9%; Score 332.5; DB 2; Length 619;  
Best Local Similarity 70.0%; Pred. No. 28-15;  
Matches 70; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVNVEDFK 60

Db 223 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVNVEDFK 282

QY 61 NSNGEAEQYRAAGEDLAQAELEKTEADLKAVHEPE 100

Db 283 ENNVVE-DYFKEGLEKTTAAKAELEKTEADLKAVHEPE 321

RESULT 2

A41971

surface protein psPA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A>Title: Structural properties and evolutionary relationships of PsPA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971

A>Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-619 <YOT>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.



C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T38077  
R/Connor, R.; Churcher, C.M.; Bartell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, April 1996  
A/Reference number: Z21767  
A/Accession: T38077  
A/Status: preliminary; translated from GE/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1957 <CON>  
A/Cross-references: UNIPROT:Q10411; EMBL:Z70690; PIDN:CAA94624.1; GSPDB:GNO0066; SPDB:S  
C/Genetics:  
A/Experimental source: strain 972h-; cosmid c1f3  
A/Gene: SPDB:SPAC1f3.06c  
A/Map position: 1

```

RESULT 9
S43074
epidermal growth factor receptor substrate - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43074; I38525
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.
Oncogene 9, 1039-1045, 1994
A:Title: A novel gene, AF-tp, fused to HRX in t(1;11)(p32;q23), is not related to AF-4,
A:Reference number: S43074; MUID:94181254; PMID:8134107

```

A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-896 <BER>  
A;Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA823105.1; PID:g470  
R;Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner  
Oncogene 9, 1591-1597, 1994  
A;Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in ev  
A;Reference number: I38525; MUID:94239734; PMID:8183552  
A;Accession: I38525  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-821,'N', 823-896 <RES>  
A;Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA521101.1; PID:g466260  
C;Genetics:  
A;Gene: GDB:EPS15, AF-1P; MLT5  
A;Cross-references: GDB:360337; OMIM:600051  
A;Map position: 1p32-1p32

```

Query Match      21.1%; Score 103.5; DB 2; Length 896;
Best Local Similarity 27.5%; Pred. No. 7.9;
Matches 28; Conservative 26; Mismatches 41; Indels 7; Gaps 3;

QY 3 EIDESDSDDVVKGLR--APLOSELDAKQAKLSKL-----EELSDKIDELDAETAKLEKNV 56
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 353 EQLKSEKEDTIKQRTSEVDQLDQVOREVNTNQLQAKQVQVELDELDEKQALEEQL 412

QY 57 EDFPNNGEAEQYRAAGEDIAAKQAELEKTEADLUKVAHE 98
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 413 KEVRKKCAEQAQLISLKRAE-LTSQESQISTVEEELAKREE 453

```

RESULT 10  
T24635

```
A;Accession: A43422
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A;Residues: 1900-1959 <HOD>
A;Experimental source: brush border
A;Note: sequence extracted from NCBI backbone (NCBIP:111947)
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
F;84-764/Domain: myosin motor domain homology <MMOT>
F;174-181/Region: nucleotide-binding motif A (P-loop)
F;552-565/Region: actin binding #status predicted
F;626-640/Region: actin binding #status predicted
F;837-1936/Domain: actin binding #status predicted <COI>
F;837-1277/Region: S2
F;1278-1959/Region: light meromyosin
F;1937-1959/Domain: carboxyl-terminal <CBT>
F;125/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F;180/Binding site: ATP (Lys) #status predicted
F;694,704/Active site: Cys #status predicted

Query Match          20.8%; Score 102; DB 1; Length 1959;
Best Local Similarity 30.3%; Pred. No. 22;
Matches 30; Conservative 24; Mismatches 31; Indels 14; Gaps 3;

QY   3 EIDSESDSYVKE-----GLRAPLOSELDAKQAKLSKLSELS-----DKIDEIDAEL 49
DB   1054 EGDSDLHQIAELQAIAELQTSLSKSEELQAALRAVEEAAQNMAKKIRELESQI 1113

QY   50 AKLEKNVEDFKNSGEEAFQYRAAGCEDLAAKQAELEKT 88
DB   1114 TELQEDLES-ERASRNKAQRKRDLGEELEALKTELED 1151

RESULT 12
T34418
Hypothetical protein F12F3.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34418
R;Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid F12F3.
A;Reference number: Z21521
A;Accession: T34418
A;Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: DNA
A;Residues: 1-3488 <FUL>
A;Cross-references: EMBL:U800022; PIDN:AC25885.1; GSPDB:GN00023; CESP:F12F3.3
A;Experimental source: strain Bristol N2; clone F12F3
C;Genetics:
A;Gene: CESP:F12F3.3
A;Map position: 5
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match          20.8%; Score 102; DB 2; Length 3488;
Best Local Similarity 32.8%; Pred. No. 38;
Matches 41; Conservative 22; Mismatches 28; Indels 34; Gaps 7;

QY   2 KEIDES---DSEYVVEGLRAPLOSELDAKQAKLSKL-----EELSDELDIDAEL 49
DB   1009 KETDEKILKDAETAATKQEADEKSKLDA-QEKIKKVSEDDAARKKELNDKL-KLESEI 1066

QY   50 A-----KLEKNVEDFKNSGEEAFQYR-----AAAGEDLAAKQAELEKTEA 90
DB   1067 ATKASADKLLREQAQAKAAEVAEAAKKQEKXDEQLKLDTEAAASKCAAEKLELEK-QA 1125

QY   91 DLKKA 95
DB   1126 QIKKA 1130

RESULT 13
C70445
```

ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus  
C;Species: Aquifex aeolicus  
C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C;Accession: C70445  
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V.  
Nature 392, 353-358, 1998  
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A;Reference number: A70300; MUID:98196666; PMID:9537320  
C;Accession: C70445  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1006 <AQF>  
A;Cross-references: UNIPROT:O67588; GB:AE000750; MID:g2983999; PIDN:AA07550.1; PID:g298  
A;Experimental source: strain VF5  
C;Genetics:  
A;Gene: clipB  
C;Superfamily: endopeptidase Clp ATP-binding chain  
C;Keywords: hydrolase

Query Match 20.4%; Score 100; DB 2; Length 1006;  
Best Local Similarity 31.7%; Pred. No. 15;  
Matches 33; Conservative 22; Mismatches 33; Indels 16; Gaps 4;  
QY 1 LKEIDSDSE-----DYVKEGLRAPLOSELDAKQAKLSK-LELSKIDELDAETAKLEK 54  
Db 552 IKALEEQITTEANLKGDYKE-----AQLKFAKLEKEKQELLGKVGVEAKIAELKK 604  
QY 55 NVEDFKNSGEEAEQYRAAGEDLAAKQAELEKTEADLKKAHVE 98  
Db 605 KIELDEKIEAKEGDYKEAELEKIEKAKLEK---ELKKLEQE 645

RESULT 14  
A28313  
glued protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C;Accession: A28313  
R;Swaroop, A.; Swaroop, M.; Garen, A.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987  
A;Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ge  
A;Reference number: A28313; MUID:87317680; PMID:2819881  
A;Accession: A28313  
A;Molecule type: DNA; mRNA  
A;Residues: 1-1319 <SWA>  
A;Cross-references: UNIPROT:PI3496  
A;Note: the authors' translation is inconsistent with the nucleotide sequence in the reg  
C;Genetics:  
A;Gene: FlyBase:Gl  
A;Cross-references: FlyBase:FBgn0001108  
A;Introns: 18/2; 479/3  
C;Keywords: cytoskeleton; glycoprotein  
F;397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 20.4%; Score 100; DB 2; Length 1319;  
Best Local Similarity 30.2%; Pred. No. 20;  
Matches 35; Conservative 18; Mismatches 25; Indels 38; Gaps 5;  
QY 1 LKEIDSDSEDYVKEGLRAPLOSELDAKQAKLSK-----ELSDKIDELDAETAKLEKRV 56  
Db 429 LRDLSAHKHDQK-----LSKELEMKRSEVTELETKELSAKIDELDAETAVADIQ---- 479  
QY 57 EDFKNSGEEAEQYRAAG-----EDLAQAELEK-----TEADLKKAHVE 98  
Db 480 -----EQVDALGAEMVEQLAEKKMELEDKVKLLBEEIAQLAELEHVE 524

RESULT 15  
H69378  
conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69378  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Usterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
C;Accession: H69378  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-886 <KLE>  
A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; MID:g2689355; PIDN:AA8902  
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 20.3%; Score 99.5; DB 2; Length 886;  
Best Local Similarity 26.0%; Pred. No. 15;  
Matches 34; Conservative 27; Mismatches 33; Indels 37; Gaps 5;  
QY 1 LKEIDSDSEDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEK----- 55  
Db 303 LRDVKEKREG-DLITREA--AGIQALKKAEDNSKLEETITKRIEELERELERFEKSHRLLE 359  
QY 56 -----VEDPKN-----SNGEEAEQYRAAGEDLAQAELEK 87  
Db 360 TLKPKMDRMQGIKAKLEE-KNLTDPKVKQWYDLLSKAKEEKEITEKLLIAKSSLKT 418  
QY 88 TEADLKKAHVE 98  
Db 419 RGAQLKKAHVE 429

Search completed: June 18, 2005, 17:03:51  
Job time : 14.113 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)

840.012 Million cell updates/sec.

Title: US-10-674-755-3

Perfect score: 490

Sequence: 1 LKEIDSESDYVKEGLRAP.....KQAELEKTEADLKAVHEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	98.0	416	2 Q9LAY8	Q9LAY8 streptococc
2	472	96.3	222	2 Q9L577	Q9L577 streptococc
3	472	96.3	262	2 Q9L576	Q9L576 streptococc
4	472	96.3	415	2 Q9LAY7	Q9LAY7 streptococc
5	463	94.5	225	2 Q9L591	Q9L591 streptococc
6	431	88.0	406	2 Q9LAZ0	Q9LAZ0 streptococc
7	430	87.8	340	2 Q8KQK5	Q8KQK5 streptococc
8	428	87.3	394	2 Q9LAY6	Q9LAY6 streptococc
9	428	87.3	395	2 Q9LAZ1	Q9LAZ1 streptococc
10	421	85.9	194	2 Q9L5B5	Q9L5B5 streptococc
11	421	85.9	218	2 Q6UEB2	Q6UEB2 streptococc
12	421	85.9	233	2 Q9L568	Q9L568 streptococc
13	421	85.9	236	2 Q9L569	Q9L569 streptococc
14	421	85.9	243	2 Q9L564	Q9L564 streptococc
15	421	85.9	243	2 Q9L567	Q9L567 streptococc
16	421	85.9	244	2 Q9L565	Q9L565 streptococc
17	421	85.9	247	2 Q9L566	Q9L566 streptococc
18	421	85.9	249	2 Q9L570	Q9L570 streptococc
19	421	85.9	254	2 Q9L563	Q9L563 streptococc
20	421	85.9	401	2 Q9LAZ2	Q9LAZ2 streptococc
21	415	84.7	246	2 Q9L578	Q9L578 streptococc
22	410	83.7	255	2 Q9L581	Q9L581 streptococc
23	410	83.7	255	2 Q9L586	Q9L586 streptococc
24	397	81.0	237	2 Q9L592	Q9L592 streptococc
25	397	81.0	395	2 Q9LAY9	Q9LAY9 streptococc
26	396	80.8	393	2 Q9LAZ3	Q9LAZ3 streptococc
27	395	80.6	207	2 Q8GNS9	Q8GNS9 streptococc
28	348.5	71.1	417	2 Q9LAY3	Q9LAY3 streptococc
29	332.5	67.9	619	2 Q54972	Q54972 streptococc
30	332.5	67.9	619	2 Q8DR10	Q8DR10 streptococc
31	331.5	67.7	739	2 Q9RQT4	Q9RQT4 streptococc

```
32 331.5 67.7 820 2 Q9RQT1 Q9rqt1 streptococc
33 331.5 67.7 929 2 Q9KK19 Q9kk19 streptococc
34 331.5 67.7 929 2 Q9ZAY5 Q9zay5 streptococc
35 325.5 66.4 415 2 Q9LAY1 Q9lay1 streptococc
36 319.5 65.2 99 2 Q8KQK4 Q8kqk4 streptococc
37 315.5 64.4 437 2 Q9LAY4 Q9lay4 streptococc
38 312.5 63.8 249 2 Q9L575 Q9l575 streptococc
39 309.5 63.2 426 2 Q9L575 Q9l575 streptococc
40 305.5 62.3 224 2 Q8GNS8 Q8gns8 streptococc
41 300.5 61.3 395 2 Q9LAY2 Q9lay2 streptococc
42 300.5 61.3 408 2 Q9LAY0 Q9lay0 streptococc
43 296 60.4 869 2 Q9KK27 Q9kk27 streptococc
44 157 32.0 246 2 Q9L5B4 Q9l5b4 streptococc
45 157 32.0 653 2 Q34097 Q34097 streptococc
```

#### ALIGNMENTS

##### RESULT 1

```
Q9LAY8 PRELIMINARY; PRT; 416 AA.
AC Q9LAY8;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8838;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071807; AAF27703.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 416 416
SQ SEQUENCE 416 AA; 45987 MW; 990C8858BC6B12C7 CRC64;
```

```
Query Match 98.0%; Score 480; DB 2; Length 416;
Best Local Similarity 98.0%; Pred. No. 4e-22;
Matches 98; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 60
Db 229 LKEIDSESDYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKNVEDPK 288

QY 61 NSNGEQAEQYRAAAGEDLAQAELEKTEADLKAVHEPE 100
Db 289 NSNGEQAEQYRAAAGEDLAQAELEKTEADLKAVHEPE 328
```

##### RESULT 2

```
Q9L577 PRELIMINARY; PRT; 222 AA.
ID Q9L577;
AC Q9L577;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=130;
RX  MEDLINE=20472698; PubMed=11015380;
RA  Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT  "Pneumococcal pspA sequence types of prevalent multiresistant
RT  pneumococcal strains in the United States and of internationally
RT  disseminated clones.";
RL  J. Clin. Microbiol. 38:3663-3669 (2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=130;
RA  Beall B.W.;
RL  Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF255550; AAF68103.1; -.
FT  NON_TER 1
FT  NON_TER 222
SQ  SEQUENCE 222 AA; 24558 MW; 6D7EB7842FE9F2A6 CRC64;

Query Match 96.3%; Score 472; DB 2; Length 222;
Best Local Similarity 97.0%; Pred. No. 7.1e-22;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 60
Db 25 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 84

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVHEPE 100
Db 85 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVNEPE 124

RESULT 3
Q9L576 PRELIMINARY; PRT; 262 AA.
AC Q9L576;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255551; AAF68104.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PRO0194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 262
SQ SEQUENCE 262 AA; 29012 MW; 32C769099466A584 CRC64;

Query Match 96.3%; Score 472; DB 2; Length 262;
Best Local Similarity 97.0%; Pred. No. 8.2e-22;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 60
Db 65 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 124
```

```

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVHEPE 100
Db 125 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVNEPE 164

RESULT 4
Q9LAY7 PRELIMINARY; PRT; 415 AA.
AC Q9LAY7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG6692;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/JAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900 (2000).
DR EMBL; AF071808; AAF27704.1; -.
FT NON_TER 415
FT NON_TER 415
SQ SEQUENCE 415 AA; 45593 MW; 41375ACBFA10FA46 CRC64;

Query Match 96.3%; Score 472; DB 2; Length 415;
Best Local Similarity 97.0%; Pred. No. 1.2e-21;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 60
Db 229 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLNVEDFK 288

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVHEPE 100
Db 289 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAHVNEPE 328

RESULT 5
Q9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI99;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI99;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254258; AAF68093.1; -.

```



```
DR InterPro; IPR009082; His_kin_homodim.
FT NON TER 1
FT NON TER 225
SQ SEQUENCE 225 AA; 24835 MW; F878A7618B72A692 CRC64;

Query Match 94.5%; Score 463; DB 2; Length 225;
Best Local Similarity 95.0%; Pred. No. 2.6e-21;
Matches 95; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 34 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 93

QY 61 NSNGEAEQYRAAGEDLAQKAELEKTEADLKKAVHEPE 100
Db 94 NSNGEAEQYRAAAEEDLAQKAELEKTEADLKKAVNEPE 133

RESULT 6
Q9LAZ0 PRELIMINARY; PRT; 406 AA.
AC Q9LAZ0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=DL6A;
RX MEDLINE=20448953; PubMed=10992499;
DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 88.0%; Score 431; DB 2; Length 406;
Best Local Similarity 89.0%; Pred. No. 4.1e-19;
Matches 89; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 213 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 272

QY 61 NSNGEAEQYRAAGEDLAQKAELEKTEADLKKAVHEPE 100
Db 273 NSDGEAQGYLAAREEDLIQKAELEQTEADLKKAVNEPE 312

RESULT 7
Q8KQK5 PRELIMINARY; PRT; 340 AA.
AC Q8KQK5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
```

```
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Sc 435/96;
RX MEDLINE=22170754; PubMed=12183557;
DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
immunization with DNA vaccines against Streptococcus pneumoniae";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1
FT NON TER 340
SQ SEQUENCE 340 AA; 38023 MW; EE07ECF00B1FBD57 CRC64;

Query Match 87.8%; Score 430; DB 2; Length 340;
Best Local Similarity 89.0%; Pred. No. 4e-19;
Matches 89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 197 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 256

QY 61 NSNGEAEQYRAAGEDLAQKAELEKTEADLKKAVHEPE 100
Db 257 NSDGEAQGYLAAREEDLIQKAELEKTEADLKKAVNEPE 296

RESULT 8
Q9LAY6 PRELIMINARY; PRT; 394 AA.
AC Q9LAY6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=L81905;
RX MEDLINE=20448953; PubMed=10992499;
DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071809; AAF27705.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 394
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C88FAA CRC64;

Query Match 87.3%; Score 428; DB 2; Length 394;
Best Local Similarity 89.0%; Pred. No. 6.1e-19;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 60
Db 213 LKEIDSDSDYVYKGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAETAKLEKNVEDFK 272

QY 61 NSNGEAEQYRAAGEDLAQKAELEKTEADLKKAVHEPE 100
Db 273 NSDGEAQGYLAAREEDLIQKAELEKTEADLKKAVNEPE 312
```

```
RESULT 9
Q9LAZ1
ID Q9LAZ1 PRELIMINARY; PRT; 395 AA.
AC Q9LAZ1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9739;
RX MEDLINE=20448953; PubMed=10992499;
DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900 (2000).
DR EMBL; AF071804; AAF27700.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECAC41DB7F95 CRC64;

Query Match 87.3%; Score 428; DB 2; Length 395;
Best Local Similarity 89.0%; Pred. No. 6.1e-19;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 60
Db 213 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 272

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 273 NSDGEQAEQYLVAAKKOLDKAKAELEKTEADLKKAVDEPE 312

RESULT 10
Q9LSB5
ID Q9LSB5 PRELIMINARY; PRT; 194 AA.
AC Q9LSB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253407; AAF67355.1; -.

RESULT 11
Q6UEB2
ID Q6UEB2 PRELIMINARY; PRT; 218 AA.
AC Q6UEB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URSP2;
RX PubMed=14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;
RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;
RT "Epitope mapping of a protective monoclonal antibody against
Pneumocystis carinii with shared reactivity to Streptococcus
pneumoniae surface antigen PspA.";
RL Infect. Immun. 72:1548-1556 (2004).
DR EMBL; AY371665; AAR20918.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 218
SQ SEQUENCE 218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;

Query Match 85.9%; Score 421; DB 2; Length 218;
Best Local Similarity 86.0%; Pred. No. 9.6e-19;
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 60
Db 27 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 86

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 87 NSDGEQAEQYLVAAKKOLDKAKAELEKTEADLKKAVDEPE 126

RESULT 12
Q9L568
ID Q9L568 PRELIMINARY; PRT; 233 AA.
AC Q9L568;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
```

```
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 194
SQ SEQUENCE 194 AA; 21116 MW; B68189FCA2B244F8 CRC64;

Query Match 85.9%; Score 421; DB 2; Length 194;
Best Local Similarity 86.0%; Pred. No. 8.6e-19;
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 60
Db 55 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 114

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 115 NSDGEQAEQYLVAAKKOLDKAKAELEKTEADLKKAVDEPE 154

RESULT 13
Q6UEB2
ID Q6UEB2 PRELIMINARY; PRT; 218 AA.
AC Q6UEB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URSP2;
RX PubMed=14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;
RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;
RT "Epitope mapping of a protective monoclonal antibody against
Pneumocystis carinii with shared reactivity to Streptococcus
pneumoniae surface antigen PspA.";
RL Infect. Immun. 72:1548-1556 (2004).
DR EMBL; AY371665; AAR20918.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 218
SQ SEQUENCE 218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;

Query Match 85.9%; Score 421; DB 2; Length 218;
Best Local Similarity 86.0%; Pred. No. 9.6e-19;
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 60
Db 27 LKEIDESSEDYVKEGLRAPLQSELDKAKQAKLSKLELSDKIDELDAIAKLEKNVEDFK 86

Qy 61 NSNGEAEQYRAAAGEDLAAKQAELEKTEADLKKAVHEPE 100
Db 87 NSDGEQAEQYLVAAKKOLDKAKAELEKTEADLKKAVDEPE 126

RESULT 14
Q9L568
ID Q9L568 PRELIMINARY; PRT; 233 AA.
AC Q9L568;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
```

RP	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RX	MEDLINE=20472698; PubMed=11015380;	
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;	
RT	"Pneumococcal pspA sequence types of prevalent multiresistant	
RT	pneumococcal strains in the United States and of internationally	
RT	disseminated clones.";	
RL	J. Clin. Microbiol. 38:3663-3669(2000).	
RP	[2]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RA	Beall B.W.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF255902; AAF70091.1; -	
DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON TER 1	
FT	NON TER 233	
SQ	SEQUENCE 233 AA; 24514 MW; D5C494019C45BFE2 CRC64;	
Query Match 85.9%; Score 421; DB 2; Length 233;		
Best Local Similarity 86.0%; Pred. No.1e-18;		
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;		
Qy	1 LKIDSDSDYKVEGRLAPQSLDQAOKLSKLELSDKIDELDAEIAKLEKVEDFK 60	
Db	:     :     :     :     :     :     :     :     :	
Db	28 LKIDSDSDYIKEGRLAPQSKLDAAKAKLSKLELSDKIDELDAEIAKLEKVEDFK 87	
Qy	61 NSNGEEAEQYRAAGEDLAQAQAELEKTEADLKKAVHEP 100	
Db	:     :     :     :     :     :     :     :	
Db	88 NSDGEAQEYIVAAKKDLDAKAELENTADLKKAVDEP 127	
RESULT 13		
Q9L569	PRELIMINARY; PRT; 236 AA.	
ID	Q9L569	
AC	Q9L569;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	PspA (Fragment).	
GN	Name=pspA;	
OS	Streptococcus pneumoniae.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1313;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=177;	
RX	MEDLINE=20472698; PubMed=11015380;	
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;	
RT	"Pneumococcal pspA sequence types of prevalent multiresistant	
RT	pneumococcal strains in the United States and of internationally	
RT	disseminated clones.";	
RL	J. Clin. Microbiol. 38:3663-3669(2000).	
RP	[2]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=177;	
RA	Beall B.W.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AP255901; AAF70091.1; -	
DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON TER 1	
FT	NON TER 236	
SQ	SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;	
Query Match 85.9%; Score 421; DB 2; Length 236;		
Best Local Similarity 86.0%; Pred. No.1e-18;		
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;		
Qy	1 LKIDSDSDYKVEGRLAPQSLDQAOKLSKLELSDKIDELDAEIAKLEKVEDFK 60	
Db	:     :     :     :     :     :     :     :	
Db	49 LKIDSDSDYIKEGRLAPQSKLDAAKAKLSKLELSDKIDELDAEIAKLEKVEDFK 108	

```

RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255903; AAF70093.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match      85.9%; Score 421; DB 2; Length 243;
Best Local Similarity 86.0%; Pred. No. 1.1e-18;
Matches 86; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRAPLQSELDAAKOAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 50 LKEIDSESDYIKELRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLEKNVEDFK 109

QY 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKKAVHEPE 100
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 110 NSDGEAEQYLVAAKKDLDAKKAELENTADLKKAVDEPE 149

```

Search completed: June 18, 2005, 17:01:34  
 Job time : 61.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-4  
Perfect score: 488  
Sequence: 1 LKIDBSDSYKGERAP.....KXAELEKAEADLKXAVDEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	473	96.9	183	2	AAW14570 Streptococ
2	473	96.9	183	7	ABW02604
3	473	96.9	8991	6	ABU08487 S. pneumo
4	456	93.4	168	7	ABW02609 L81905c p
5	447	91.6	194	2	AAW14584 Streptococ
6	447	91.6	194	7	ABW02618
7	440.5	90.3	167	2	AAW14575
8	430	88.1	550	8	ADK48356 Streptococ
9	430	88.1	550	8	ADR95223 Novel S.
10	405	83.0	166	2	AAW14568 Streptococ
11	405	83.0	166	7	ABW02602
12	384.5	78.8	185	2	AAW14566
13	384.5	78.8	185	7	ABW02600
14	327.5	67.1	204	2	AAW14571 Streptococ
15	327.5	67.1	204	7	ABW02605
16	325.5	66.7	170	7	ABW02614 Rct135c p
17	325.5	66.7	181	7	ABW02596
18	325.5	66.7	865	6	ABU08489 S. pneumo
19	325.5	66.7	929	2	AAW14593 Streptococ
20	325.5	66.7	929	2	AAW14384 S. pneumo
21	322.5	66.1	188	2	AAW14580 Streptococ
22	322.5	66.1	188	7	ABW02613 Rct129c p
23	314.5	64.4	198	2	AAW14581
24	313.5	64.2	588	6	ABU08491 Coiled co
25	313.5	64.2	589	2	AAW43392 PspC alph

26	311.5	63.8	198	7	ABW02615	Abw02615 Rx1c pneu
27	311.5	63.8	204	2	AAW14578	Aaw14578 Streptococ
28	311.5	63.8	204	7	ABW02612	Abw02612 Rct123c p
29	311.5	63.8	315	2	AAV04375	Aay04375 Streptococ
30	311.5	63.8	619	2	AAW63437	Aar63437 Pneumococ
31	311.5	63.8	619	2	AAW87598	Aar87598 Pneumococ
32	311.5	63.8	619	2	AAW86911	Aar86911 Streptococ
33	311.5	63.8	619	2	AAW41838	Aay41838 Streptococ
34	311.5	63.8	619	5	AAE18782	Aae18782 S. pneumo
35	311.5	63.8	619	6	ABU45778	Abu45778 Protein e
36	311.5	63.8	619	8	ADO52126	Ado52126 Streptococ
37	311.5	63.8	648	2	AAW70336	Aaw70336 Pneumococ
38	311.5	63.8	648	2	AAW62274	Aaw62274 Streptococ
39	311.5	63.8	648	2	AAW41837	Aay41837 Streptococ
40	311.5	63.8	648	2	AAW87879	Aaw87879 A. pneumoc
41	311.5	63.8	653	2	AAW92456	Aaw92456 S. pneumo
42	311.5	63.8	684	2	AAW73912	Aar73912 Streptococ
43	311	63.7	180	2	AAW14562	Aaw14562 Streptococ
44	308	63.1	187	2	AAW14579	Aaw14579 Streptococ
45	305.5	62.6	195	2	AAW14591	Aaw14591 Streptococ

## ALIGNMENTS

### RESULT 1

AAW14570

ID AAW14570 standard; protein; 183 AA.

AC AAW14570;

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

DE Streptococcus pneumoniae PspA central region.

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

OS Streptococcus pneumoniae; strain Bg9739.

PN WO9709994-Al.

PD 20-MAR-1997.

PF 16-SEP-1996; 96WO-US014819.

PR 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

WPI; 1997-202002/18.

DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739. Comparison of the N-terminal and central regions (AAW14533-57 and AAW14562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in vaccines to protect animals against S. pneumoniae infection and hence for the prevention of diseases such as otitis media, meningitis, bacteraemia and pneumonia. The sequence of the 3' half of the PspA alpha-helical region and the immediate 5' tip of the coding sequence are likely to be the critical sequences for predicting PspA cross-reactions and vaccine

```
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 183 AA;
    Query Match          96.9%; Score 473; DB 2; Length 183;
    Best Local Similarity 97.0%; Pred. No. 2e-35;
    Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60
Db 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60

Qy 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAVIDEPE 100
Db 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAVIDEPE 100

RESULT 2
ABW02604
ID ABW02604 standard; protein; 183 AA.
XX
AC ABW02604;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bg9739c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
PT Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 50; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bg9739c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 183 AA;
    Query Match          96.9%; Score 473; DB 7; Length 183;
    Best Local Similarity 97.0%; Pred. No. 2e-35;

Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKVEDFK 60
Db 5139 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKVEDFK 5198
```

QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVDEPE 100  
Db 5199 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVDEPE 5238

## RESULT 4

ABW02609  
ID ABW02609 standard; protein; 168 AA.

XX ABW02609;  
XX AC  
XX DT  
XX DT  
XX DT  
XX DE L81905c pneumococcal surface protein A (PspA) central region.  
XX PspA; pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX immunological; gene therapy; immunostimulant.  
XX KW  
XX OS Unidentified.  
XX

XX Key Location/Qualifiers  
FH Misc-difference 1.168  
FT /note= "Xaa = Unknown amino acid"  
XX

US6592876-B1.  
XX

15-JUL-2003.  
XX

15-SEP-1995; 95US-00529055.  
XX

20-APR-1993; 93US-00048896.  
XX

06-JUN-1995; 95US-00465746.  
XX

(UABR-) UAB RES FOUND.  
XX

Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX

WPI; 2003-862841/80.  
XX

XX Immunological composition for obtaining expression products used for  
XX detecting the presence of Streptococcus pneumoniae or its strain.  
XX comprises at least two different full length isolated gene encoding  
XX pneumococcal surface protein A.

XX Example 6; SEQ ID NO 55; 121pp; English.  
XX

XX The present invention relates to an immunological composition comprising  
XX at least 2 different full length isolated genes encoding pneumococcal  
XX surface protein A (PspA) from different groups based on restriction  
XX fragment polymorphism analysis. The invention is useful for obtaining  
XX expression products by recombinant techniques to detect, determine,  
XX isolate or diagnose the presence of Streptococcus pneumoniae or its  
XX strain. The expression product is useful for preparing antigenic,  
XX immunological or vaccine compositions, for eliciting antibodies, an  
XX immunological response (other than or additional to antibodies) or a  
XX protective response (including antibody or other immunological response  
XX by administering compositions to a host). The invention is also useful as  
XX vaccines and in gene therapy. The present sequence is L81905c  
XX pneumococcal surface protein A (PspA) central region. This sequence is  
XX used in the exemplification of the invention

SQ Sequence 168 AA;

Query Match 93.4%; Score 456; DB 7; Length 168;  
Best Local Similarity 95.0%; Pred. No. 6.4e-34;  
Matches 95; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEDISSDSEDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60  
Db 1 LKEDISSDSEDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVDEPE 100  
|||||

Db 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVDEPE 100

## RESULT 5

AAW14584  
ID AAW14584 standard; protein; 194 AA.

XX AAW14584;  
XX AC

XX 17-OCT-2003 (revised)  
XX DT

XX 28-OCT-1997 (first entry)  
XX DT

XX Streptococcus pneumoniae PspA central region.  
XX DE

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
XX KW bacteriaemia; pneumonia.

XX Streptococcus pneumoniae; strain Db16.  
XX OS

XX Key Location/Qualifiers  
FH Misc-difference 61  
FT /note= "unidentified amino acid"

XX WO9709994-A1.  
XX PN

20-MAR-1997.  
XX PD

16-SEP-1996; 96WO-US014819.  
XX PF

15-SEP-1995; 95US-00529055.  
XX PR

(UABR-) UAB RES FOUND.  
XX PA

Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
XX PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.  
XX DR

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
XX in vaccines for protecting animals against S.pneumoniae infection.  
XX Example 6; Fig 13; 296pp; English.  
XX

XX This sequence shows the central portion, including the C-terminus of the  
XX alpha-helix region and some of the proline-rich region, of pneumococcal  
XX surface protein A (PspA) of Streptococcus pneumoniae strain Db16.  
XX Comparison of the N-terminal and central regions (AAW14533-57 and  
XX AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
XX be used to divide the strains into several families based on sequence  
XX homologies. PspA polypeptides, or fragments of them, can be used in  
XX vaccines to protect animals against S. pneumoniae infection and hence for  
XX the prevention of diseases such as otitis media, meningitis, bacteriaemia  
XX and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
XX region and the immediate 5' tip of the coding sequence are likely to be  
XX the critical sequences for predicting PspA cross-reactions and vaccine  
XX composition. (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 194 AA;

Query Match 91.6%; Score 447; DB 2; Length 194;  
Best Local Similarity 93.9%; Pred. No. 5e-33;  
Matches 93; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEDISSDSEDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60  
Db 1 LKEDISSDSEDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60

QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVDEPE 99  
|||||

Db 61 XSDGEQAGQYLAAGEDLIAKAELEKAEADLKAVNEP 99  
|||||

RESULT 6





XX AC ADK48356;  
XX DT 20-MAY-2004 (first entry)  
XX DE Streptococcus pneumoniae protein, Seq ID No 4871.  
XX KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
XX OS Streptococcus pneumoniae.  
XX PN US6699703-B1.  
XX PD 02-MAR-2004.  
XX PF 26-MAY-2000; 2000US-00583110.  
XX PR 02-JUL-1997; 97US-0051553P.  
XX PR 12-MAY-1998; 98US-0085131P.  
XX PR 30-JUN-1998; 98US-00107433.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewartz CE;  
XX WPI; 2004-212399/20.  
XX DR N-PSDB; ADK45695.  
XX CC New nucleic acid molecules and polypeptides useful for diagnosing,  
PT preventing and treating pathological conditions resulting from bacterial  
PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
PT screening.  
XX PS Disclosure; SEQ ID NO 4871; 301pp; English.  
XX CC The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX SQ Sequence 550 AA;  
Query Match 88.1%; Score 430; DB 8; Length 550;  
Best Local Similarity 90.0%; Pred. No. 6.1e-31;  
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLEELSDKIDELDAETAKLEKVEDFK 60  
Db 144 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLEELSDKIDELDAETAKLEKVEDFK 203  
QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVIDEPE 100  
Db 204 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVIDEPE 243  
RESULT 9  
ADR95223  
ID ADR95223 standard; protein; 550 AA.  
XX ADR95223;  
XX AC ADR95223;  
XX DT 16-DEC-2004 (first entry)  
XX DE Novel S. pneumoniae protein sequence, SEQ ID 3858.  
XX KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;  
XX bacterial infection.

OS Streptococcus pneumoniae.  
XX US6800744-B1.  
XX PD 05-OCT-2004.  
XX PF 30-JUN-1998; 98US-00107433.  
XX PR 02-JUL-1997; 97US-0051553P.  
XX PR 12-MAY-1998; 98US-0085131P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Doucette-Stamm LA, Bush D;  
XX WPI; 2004-697205/68.  
XX DR N-PSDB; ADR92620.  
XX CC New isolated nucleic acid encoding a Streptococcus pneumoniae  
PT polypeptide, useful for diagnosing, preventing and/or treating  
PT pathological conditions resulting from the bacterial infection.  
XX PS Disclosure; SEQ ID NO 3858; 151pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising a sequence  
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its  
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,  
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,  
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,  
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or  
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide  
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is  
CC hybridisable under high stringency conditions to the nucleotide sequence.  
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.  
CC Also included are a recombinant expression vector comprising the isolated  
CC nucleic acid cited above operably linked to a transcription regulatory  
CC element, a cell comprising the recombinant expression vector and a probe  
CC comprising at least 20 consecutive nucleotides of the nucleotide  
CC sequences as cited above. The methods and compositions of the present  
CC invention are useful for the diagnosis, prevention and/or treatment of  
CC pathological conditions resulting from bacterial infection by  
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and  
CC otitis media. The present sequence is one of the 2603 disclosed S.  
CC pneumoniae protein sequences. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.  
XX SQ Sequence 550 AA;  
Query Match 88.1%; Score 430; DB 8; Length 550;  
Best Local Similarity 90.0%; Pred. No. 6.1e-31;  
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLEELSDKIDELDAETAKLEKVEDFK 60  
Db 144 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLEELSDKIDELDAETAKLEKVEDFK 203  
QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVIDEPE 100  
Db 204 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVIDEPE 243  
RESULT 10  
AAW14568  
ID AAW14568 standard; protein; 166 AA.  
XX AAW14568;  
XX AC AAW14568;  
XX DT 17-OCT-2003 (revised)  
XX DT 28-OCT-1997 (first entry)  
XX DE Streptococcus pneumoniae PepA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Bg8743.  
 OS WO9709994-A1.  
 XX  
 XX 20-MAR-1997.  
 PD  
 XX 16-SEP-1996; 96WO-US014819.  
 PF  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT  
 XX Example 6; Fig 13; 296pp; English.  
 PS  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 166 AA;  
 SQ  
 Query Match 83.0%; Score 405; DB 2; Length 166;  
 Best Local Similarity 84.0%; Pred. No. 2.9e-29;  
 Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 DB 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 QY 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAADPE 100  
 DB 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAADPE 100  
 RESULT 11  
 AEW02602  
 ID ABW02602 standard; protein; 166 AA.  
 XX  
 AC ABW02602;  
 XX  
 XX 12-FEB-2004 (first entry)  
 DT  
 XX Bg8743c pneumococcal surface protein A (PspA) central region.  
 DE  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 OS  
 XX US6592876-B1.  
 PN  
 XX 15-JUL-2003.  
 PD

XX 15-SEP-1995; 95US-00529055.  
 PF  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 XX 06-JUN-1995; 95US-00465746.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 48; 121pp; English.  
 PS  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Bg8743c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 XX Sequence 166 AA;  
 SQ  
 Query Match 83.0%; Score 405; DB 7; Length 166;  
 Best Local Similarity 84.0%; Pred. No. 2.9e-29;  
 Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 DB 1 LKEIDESDSEYVKEGERAPLQSELDKAKQKLSKLELSKIDELDAEIAKLEKDVDFK 60  
 QY 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAADPE 100  
 DB 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAADPE 100  
 RESULT 12  
 AAW14566  
 ID AAW14566 standard; protein; 185 AA.  
 XX  
 AC AAW14566;  
 XX  
 XX 17-OCT-2003 (revised)  
 DT  
 XX 28-OCT-1997 (first entry)  
 DT  
 XX Streptococcus pneumoniae PspA central region.  
 DE  
 XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 XX Streptococcus pneumoniae; strain Ac94.  
 OS  
 XX WO9709994-A1.  
 PN  
 XX 20-MAR-1997.  
 PD  
 XX 16-SEP-1996; 96WO-US014819.  
 PF  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 78.8%; Score 384.5; DB 2; Length 185;  
 Best Local Similarity 81.2%; Pred. No. 2.5e-27;  
 Matches 82; Conservative 6; Mismatches 12; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYVYKEGERAPLQSELDKAKLKLSELSKDIDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDSDSDYVYKEGERAPLQSELDVQAKLKLSELSKDIDELDAETAKLNKDVDF 60  
 QY 60 KNSDGEQAGYLAAGEDLIKAKAELEKAEADLKKAVIDEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVIDEPE 101  
 RESULT 13  
 ID ABW02600  
 AC ABW02600;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 XX  
 PT Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,

PT comprises at least two different full length isolated gene encoding  
 XX pneumococcal surface protein A.  
 PS Example 6; SEQ ID NO 46; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 78.8%; Score 384.5; DB 7; Length 185;  
 Best Local Similarity 81.2%; Pred. No. 2.5e-27;  
 Matches 82; Conservative 6; Mismatches 12; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYVYKEGERAPLQSELDKAKLKLSELSKDIDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDSDSDYVYKEGERAPLQSELDVQAKLKLSELSKDIDELDAETAKLNKDVDF 60  
 QY 60 KNSDGEQAGYLAAGEDLIKAKAELEKAEADLKKAVIDEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVIDEPE 101  
 RESULT 14  
 ID AAW14571  
 AC AAW14571;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Ef1019.  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.  
CC Comparison of the N-terminal and central regions (AAM14533-57 and  
CC AAM14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field).

XX SQ Sequence 204 AA;

Query Match 67.1%; Score 327.5; DB 2; Length 204;  
Best Local Similarity 68.8%; Pred. No. 4.6e-22;  
Matches 75; Conservative 5; Mismatches 10; Indels 19; Gaps 2;  
Qy 1 LKEIDESDSYVKEGFRAPLQSELDKAKQAKLSKLELSDKIDELDAEIAKLE----- 53  
Db 1 LKEIDESDSYVKEGFRAPLQSELDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAAE 60  
Qy 54 --KDVEDPKNSDGEQAGQYLAAGAGDILIAKKAELKAEADLKKAVDPE 100  
Db 61 ENNVEDYFKEGLEKT-----IAAKKAELKTEADLKKAVDPE 99

RESULT 15  
ABW02605  
ID ABW02605 standard; protein; 204 AA.  
AC ABW02605;  
XX  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Ef1019c pneumococcal surface protein A (PspA) central region.  
XX  
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
XX Unidentified.  
XX  
XX US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
XX WPI; 2003-862841/80.

XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 51; 121pp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspA) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Ef1019c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention

XX SQ Sequence 204 AA;

Query Match 67.1%; Score 327.5; DB 7; Length 204;  
Best Local Similarity 68.8%; Pred. No. 4.6e-22;  
Matches 75; Conservative 5; Mismatches 10; Indels 19; Gaps 2;  
Qy 1 LKEIDESDSYVKEGFRAPLQSELDKAKQAKLSKLELSDKIDELDAEIAKLE----- 53  
Db 1 LKEIDESDSYVKEGFRAPLQSELDKAKQAKLSKLELSDKIDELDAEIAKLEDLQKAAE 60  
Qy 54 --KDVEDPKNSDGEQAGQYLAAGAGDILIAKKAELKAEADLKKAVDPE 100  
Db 61 ENNVEDYFKEGLEKT-----IAAKKAELKTEADLKKAVDPE 99

Search completed: June 18, 2005, 16:51:20  
Job time : 74.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-4  
Perfect score: 488  
Sequence: 1 LKEIDSESDYVKEGERAP.....KKALEKAEADLKKAVDPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PTUS COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	100.0	100	4	US-09-147-875A-4
2	477.5	97.8	101	2	US-08-710-749-3
3	473	96.9	183	4	US-08-529-055-50
4	473	96.9	8991	4	US-08-714-741-32
5	465	95.3	100	4	US-09-147-875A-5
6	464	95.1	100	4	US-09-147-875A-6
7	456.5	93.5	101	2	US-08-710-743-4
8	456	93.4	168	4	US-08-529-055-55
9	453.5	92.9	101	2	US-08-710-749-5
10	447	91.6	194	4	US-08-529-055-64
11	433	88.7	100	4	US-09-147-875A-3
12	430	88.1	98	4	US-09-147-875A-1
13	430	88.1	550	4	US-09-583-110-4871
14	430	88.1	550	4	US-09-107-433-3858
15	428	87.7	100	4	US-09-147-875A-2
16	427.5	87.6	101	2	US-08-710-749-2
17	421.5	86.4	99	2	US-08-710-749-9
18	419.5	86.0	101	2	US-08-710-749-1
19	405	83.0	166	4	US-08-529-055-48
20	392.5	80.4	101	2	US-08-710-749-6
21	390	79.9	100	4	US-09-147-875A-7
22	384.5	78.8	185	4	US-08-529-055-46
23	382.5	78.4	101	4	US-09-147-875A-9
24	376	77.0	100	4	US-09-147-875A-8
25	374	76.6	102	2	US-08-710-743-8
26	372.5	76.3	101	2	US-08-710-749-7
27	333.5	68.3	99	4	US-09-147-875A-16

28	327.5	67.1	99	2	US-08-710-749-10	Sequence 10, Appl
29	327.5	67.1	99	4	US-09-147-875A-11	Sequence 11, Appl
30	327.5	67.1	204	4	US-08-529-055-51	Sequence 51, Appl
31	325.5	66.7	170	4	US-08-529-055-60	Sequence 60, Appl
32	325.5	66.7	181	4	US-08-529-055-42	Sequence 42, Appl
33	325.5	66.7	864	4	US-08-714-741-40	Sequence 40, Appl
34	324.5	66.5	99	2	US-08-710-749-17	Sequence 17, Appl
35	322.5	66.1	99	2	US-08-710-749-15	Sequence 15, Appl
36	322.5	66.1	188	4	US-08-529-055-59	Sequence 59, Appl
37	322	66.0	100	4	US-09-147-875A-10	Sequence 10, Appl
38	318	65.2	100	4	US-09-147-875A-12	Sequence 12, Appl
39	313.5	64.2	141	4	US-09-286-981B-2	Sequence 2, Appl
40	313.5	64.2	588	4	US-08-714-741-42	Sequence 42, Appl
41	311.5	63.8	99	2	US-08-710-749-11	Sequence 11, Appl
42	311.5	63.8	198	4	US-08-529-055-61	Sequence 61, Appl
43	311.5	63.8	204	4	US-08-529-055-58	Sequence 58, Appl
44	311.5	63.8	619	1	US-08-465-746-2	Sequence 2, Appl
45	311.5	63.8	619	1	US-08-214-164-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-147-875A-4  
; Sequence 4, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-4

Query Match 100.0%; Score 488; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.2e-39;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPLQSELDKAKLSEELSDKIDELDAETAKLEKQVDFK 60  
DB 1 LKEIDSESDYVKEGERAPLQSELDKAKLSEELSDKIDELDAETAKLEKQVDFK 60

QY 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAVDPE 100  
DB 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAVDPE 100

RESULT 2  
US-08-710-749-3  
; Sequence 3, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

RESULT 3  
US-08-529-055-50  
; Sequence 50, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:

```

, ZAP: 10038
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/714,741
, FILING DATE: 16-SEP-1996
, CLASSIFICATION: 435
, ATTORNEY/AGENT INFORMATION:
, NAME: Frommer Esq., William S.
, REGISTRATION NUMBER: 25,506
, REFERENCE/DOCKET NUMBER: 454312-2460
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (212) 840-3333
, TELEFAX: (212) 840-0712
, INFORMATION FOR SEQ ID NO: 32:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 8991 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: amino acid
, US-08-714-741-32

```

Query Match 96.9%; Score 473; DB 4; Length 8991;  
Best Local Similarity 97.0%; Pred. No. 1.3e-35;  
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60  
DB 5139 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100  
DB 5199 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 5238

RESULT 5  
US-09-147-875A-5  
; Sequence 5, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-5

Query Match 95.3%; Score 465; DB 4; Length 100;  
Best Local Similarity 96.0%; Pred. No. 3.4e-37;  
Matches 96; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60  
DB 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100  
DB 61 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100

RESULT 6  
US-09-147-875A-6  
; Sequence 6, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-09-147-875A-6

Query Match 95.1%; Score 464; DB 4; Length 100;  
Best Local Similarity 96.0%; Pred. No. 4.3e-37;  
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60  
DB 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100  
DB 61 NSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100

RESULT 7  
US-08-710-749-4  
; Sequence 4, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-4

Query Match 93.5%; Score 456.5; DB 2; Length 101;  
Best Local Similarity 95.0%; Pred. No. 2.2e-36;  
Matches 96; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDF 59  
DB 1 LKEIDSDSDYVKEGERAPLQSELDKAKLSKLEELSDKIDELDAEIAKLEKDVDF 60

QY 60 KNSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 100  
DB 61 KNSDGEQAGQYLAAGBEDIKAKAELEKAEADLKKAVIDEPE 101

RESULT 8  
US-08-529-055-55  
; Sequence 55, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions

;; TITLE OF INVENTION: Thereof, Expression Products  
;; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
;; TITLE OF INVENTION: Portions and Products  
;; NUMBER OF SEQUENCES: 73  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Curtis, Morris & Safford, P.C.  
;; STREET: 530 Fifth Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/529,055  
;; FILING DATE: 15-SEP-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2400  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 55:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 168 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-529-055-55

Query Match 93.4%; Score 456; DB 4; Length 168;  
Best Local Similarity 95.0%; Pred. No. 4.6e-36;  
Matches 95; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60  
Db 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60  
Qy 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDPE 100  
Db 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDPE 100

RESULT 9  
US-08-710-749-5  
; Sequence 5, Application US/08/10749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/710,749  
;; FILING DATE: 20-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 101 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
US-08-710-749-5  
Query Match 92.9%; Score 453.5; DB 2; Length 101;  
Best Local Similarity 95.0%; Pred. No. 4.3e-36;  
Matches 96; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
Qy 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 59  
Db 1 LKEIDSDSDYVKEGERAPLQSELDDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 60  
Qy 60 KNSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDPE 100  
Db 61 KNSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVDPE 101  
RESULT 10  
US-08-529-055-64  
; Sequence 64, Application US/08/529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 194 amino acids



```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64

Query Match      91.6%; Score 447; DB 4; Length 194;
Best Local Similarity 93.9%; Pred. No. 3.9e-35;
Matches 93; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60
DB 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60

QY 61 NSDGEQAGQYLAAGEDLIKAKAELEKAEADLKAVDEP 99
DB 61 XSDGEQAGQYLAAGEDLIKAKAELEKAELEQTEADLKAVNEP 99

RESULT 11
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

Query Match      88.7%; Score 433; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 3.8e-34;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60
DB 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60

QY 61 NSDGEQAGQYLAAGEDLIKAKAELEKAEADLKAVDEP 100
DB 61 NSNGEAEQYRAAGEDLAKQAELEKTEADLKAVNEP 100

RESULT 12
US-09-147-875A-1
; Sequence 1, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-1

Query Match      88.1%; Score 430; DB 4; Length 98;
Best Local Similarity 94.0%; Pred. No. 7.1e-34;
Matches 94; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60
```

```

DB 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60
QY 61 NSDGEQAGQYLAAGEDLIKAKAELEKAEADLKAVDEP 100
DB 61 NSDGEQAGQYLAAGEDLIKAKAELEKAEADLKAVNEP 98

RESULT 13
US-09-583-110-4871
; Sequence 4871, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4871
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4871

Query Match      88.1%; Score 430; DB 4; Length 550;
Best Local Similarity 90.0%; Pred. No. 5.6e-33;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60
DB 144 LKEIDESDSEDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 203
QY 61 NSDGEQAGQYLAAGEDLIKAKAELEKAEADLKAVDEP 100
DB 204 NSNGEAEQYRAAGEDLAKQAELEKTEADLKAVNEP 243

RESULT 14
US-09-107-433-3858
; Sequence 3858, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085131
; FILING DATE: May 12, 1998
```

```

; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinitello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match      88.1%; Score 430; DB 4; Length 550;
Best Local Similarity 90.0%; Pred. No. 5.6e-33; Mismatches 6; Indels 0; Gaps 0;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPQSELDAAKQAELEKAEADLKKAVDEPE 60
   |||||
Db 144 LKEIDSESDYVKEGLRAPQSELDAAKQAELEKAEADLKKAVDEPE 203
   |||||

QY 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVDEPE 100
   |||||
Db 204 NSNGEQAEQYRAAEEDLAAKQAELEKAEADLKKAVDEPE 243
   |||||

RESULT 15
US-09-147-875A-2
; Sequence 2, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-2

Query Match      87.7%; Score 428; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 1.1e-33; Mismatches 7; Indels 0; Gaps 0;
Matches 90; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPQSELDAAKQAELEKAEADLKKAVDEPE 60
   |||||
Db 1 LKEIDSESDYVKEGLRAPQSELDAAKQAELEKAEADLKKAVDEPE 60
   |||||

QY 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVDEPE 100
   |||||
Db 61 NSNGEQAEQYRAAEEDLAAKQAELEKAEADLKKAVDEPE 100
   |||||
```

Search completed: June 18, 2005, 17:07:05  
Job time : 18.9189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-4  
Perfect score: 488  
Sequence: 1 LKEDSDSDYKGERAP.....KKALEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 171042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 171042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	100.0	100	15 US-10-674-755-4	Sequence 4, Appli
2	473	96.9	183	15 US-10-299-636-65	Sequence 65, Appl
3	465	95.3	100	15 US-10-674-755-5	Sequence 5, Appli
4	464	95.1	100	15 US-10-674-755-6	Sequence 6, Appli
5	456	93.4	168	15 US-10-299-636-70	Sequence 79, Appl
6	447	91.6	194	15 US-10-299-636-79	Sequence 70, Appl
7	433	88.7	100	15 US-10-674-755-3	Sequence 3, Appli
8	430	88.1	98	15 US-10-674-755-1	Sequence 1, Appli
9	428	87.7	100	15 US-10-674-755-2	Sequence 2, Appli
10	405	83.0	166	15 US-10-299-636-63	Sequence 63, Appli
11	390	79.9	100	15 US-10-674-755-7	Sequence 7, Appli

12	384.5	78.8	185	15	US-10-299-636-61	Sequence 61, Appl
13	382.5	78.4	101	15	US-10-674-755-9	Sequence 9, Appli
14	376	77.0	100	15	US-10-674-755-8	Sequence 8, Appli
15	333.5	68.3	99	15	US-10-674-755-16	Sequence 16, Appl
16	327.5	67.1	99	15	US-10-674-755-11	Sequence 11, Appl
17	327.5	67.1	204	15	US-10-299-636-66	Sequence 66, Appl
18	325.5	66.7	170	15	US-10-299-636-75	Sequence 75, Appl
19	325.5	66.7	181	15	US-10-299-636-57	Sequence 57, Appl
20	325.5	66.7	643	15	US-10-299-636-95	Sequence 95, Appl
21	325.5	66.7	670	9	US-09-748-875-63	Sequence 63, Appl
22	325.5	66.7	670	10	US-09-298-523B-63	Sequence 63, Appl
23	325.5	66.7	690	9	US-09-748-875-61	Sequence 61, Appl
24	325.5	66.7	690	10	US-09-298-523B-61	Sequence 61, Appl
25	325.5	66.7	691	9	US-09-748-875-1	Sequence 1, Appli
26	325.5	66.7	691	10	US-09-298-523B-1	Sequence 1, Appli
27	325.5	66.7	701	9	US-09-748-875-62	Sequence 62, Appl
28	325.5	66.7	701	10	US-09-298-523B-62	Sequence 62, Appl
29	325.5	66.7	707	9	US-09-748-875-2	Sequence 2, Appli
30	325.5	66.7	707	10	US-09-298-523B-2	Sequence 2, Appli
31	325.5	66.7	711	9	US-09-748-875-3	Sequence 3, Appli
32	325.5	66.7	711	10	US-09-298-523B-3	Sequence 3, Appli
33	325.5	66.7	739	17	US-10-732-923-3294	Sequence 3294, Ap
34	325.5	66.7	929	9	US-09-748-875-60	Sequence 60, Appl
35	325.5	66.7	929	10	US-09-298-523B-60	Sequence 60, Appl
36	325.5	66.7	929	15	US-10-299-636-94	Sequence 94, Appl
37	322.5	66.1	188	15	US-10-674-755-10	Sequence 74, Appl
38	322.5	66.0	100	15	US-10-674-755-10	Sequence 10, Appl
39	318	65.2	100	15	US-10-674-755-12	Sequence 12, Appl
40	313.5	64.2	141	14	US-10-254-995-2	Sequence 2, Appli
41	313.5	64.2	589	9	US-09-748-875-14	Sequence 14, Appl
42	313.5	64.2	589	10	US-09-298-523B-14	Sequence 14, Appl
43	313.5	64.2	589	15	US-10-299-636-97	Sequence 97, Appl
44	311.5	63.8	198	15	US-10-299-636-76	Sequence 76, Appl
45	311.5	63.8	204	15	US-10-299-636-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1  
US-10-674-755-4  
; Sequence 4, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-4

Query Match 100.0%; Score 488; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 9.9e-33;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEDSDSDYKGERAPLOSDAKQAKLSKLELSKIDELDAETAKLEKQVEDFK 60  
Db 1 LKEDSDSDYKGERAPLOSDAKQAKLSKLELSKIDELDAETAKLEKQVEDFK 60  
QY 61 NSDGEQAGQYLAAGGEDLIKAELEKAEADLKKAVIDEPE 100  
Db 61 NSDGEQAGQYLAAGGEDLIKAELEKAEADLKKAVIDEPE 100

RESULT 2

```

US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match          96.9%; Score 473; DB 15; Length 183;
Best Local Similarity 97.0%; Pred. No. 3.3e-31;
Matches 97; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
DB 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100
DB 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100

RESULT 3
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match          95.3%; Score 465; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 7.6e-31;
Matches 96; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
DB 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100
DB 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100

```

```

RESULT 4
US-10-674-755-6
; Sequence 6, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match          95.1%; Score 464; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 9.2e-31;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
DB 1 LKEIDESSEDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

QY 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100
DB 61 NSDGEQAGQYLAAGEDIILAKKAELEKAEADLKKAVDEPE 100

RESULT 5
US-10-299-636-70
; Sequence 70, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa at position 38 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa at position 42 is unknown

```

```
;
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70

Query Match          93.4%; Score 456; DB 15; Length 168;
Best Local Similarity 95.0%; Pred. No. 7.4e-30;
Matches 93; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   |||||
Db 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   |||||

RESULT 6
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match          91.6%; Score 447; DB 15; Length 194;
Best Local Similarity 93.9%; Pred. No. 4.7e-29;
Matches 93; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEP 99
   |||||
Db 61 XSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVNEP 99
   |||||

RESULT 7
US-10-674-755-3
; Sequence 3, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
```

```
;
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match          88.7%; Score 433; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 3.2e-28;
Matches 90; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   |||||
Db 61 NSNGEAEQYRAAGEDLIAKKAELKAEADLKKAVHEPE 100
   |||||

RESULT 8
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match          88.1%; Score 430; DB 15; Length 98;
Best Local Similarity 94.0%; Pred. No. 5.5e-28;
Matches 94; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGLRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   |||||
Db 61 NSDGEQA-QYLAAGEDL-AKKAELKAEADLKKAVHEPE 98
   |||||

RESULT 9
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
```

```
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match      87.7%; Score 428; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 8.2e-28;
Matches 90; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEAQGYLAAGEDIKAELEKAEADLKKAADLKKAVDEPE 100
Db 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 10
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      83.0%; Score 405; DB 15; Length 166;
Best Local Similarity 84.0%; Pred. No. 1.1e-25;
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEAQGYLAAGEDIKAELEKAEADLKKAADLKKAVDEPE 100
Db 61 NSDGEAQGYLVAAEKDLDAKEAEELGNTGADLKKAVDEPE 100

RESULT 11
US-10-674-755-7
; Sequence 7, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-7

Query Match      79.9%; Score 390; DB 15; Length 100;
Best Local Similarity 82.0%; Pred. No. 1.1e-24;
Matches 82; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEAQGYLAAGEDIKAELEKAEADLKKAADLKKAVDEPE 100
Db 61 NSDGEAQGYLVAAEKDLDAKEAEELGNTGADLKKAVDEPE 100

RESULT 12
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-61

Query Match      78.8%; Score 384.5; DB 15; Length 185;
Best Local Similarity 81.2%; Pred. No. 6e-24;
Matches 82; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGERAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 59
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60

Qy 60 KNSDGEAQGYLAAGEDIKAELEKAEADLKKAADLKKAVDEPE 100
Db 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101

RESULT 13
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
```

```
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      78.4%; Score 382.5; DB 15; Length 101;
Best Local Similarity 81.2%; Pred. No. 4.4e-24;
Matches 82; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAK-LEKDVDF 59
   |||||
Db 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKNKKDVDF 60
   |||||

QY 60 KNSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   :|||
Db 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
   :|||

RESULT 14
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-8

Query Match      77.0%; Score 376; DB 15; Length 100;
Best Local Similarity 81.0%; Pred. No. 1.5e-23;
Matches 81; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKDVDF 60
   |||||
Db 1 LKGIIDSDSDYVKEGERAPLOSELDAKRTKLTLEELSDKIDELDAETIPKLEKNVEYFK 60
   :|||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   :|||
Db 61 LTDAEQTEQYLAAGAKDLADKKAELKTEADLKKAVHEPE 100
   :|||

RESULT 15
US-10-674-755-16
; Sequence 16, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-16

Query Match      68.3%; Score 333.5; DB 15; Length 99;
Best Local Similarity 73.3%; Pred. No. 4.5e-20;
Matches 74; Conservative 9; Mismatches 15; Indels 3; Gaps 2;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKDVDF 60
   |||||
Db 1 LKEIDSDSDYVKEGERAPLOSKLDTRKAKLSKLEELSDKIDELDAETAKLEQVKDAE 60
   :|||

QY 61 NSDGEQAGQYLAAGEDLIAKKAELKAEADLKKAVDEPE 100
   :|||
Db 61 GNNVVEA--YFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
   :|||
```

Search completed: June 18, 2005, 18:00:22  
Job time : 63.963 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-4  
Perfect score: 488  
Sequence: 1 LKEIDSESDYKGERAP.....KKAELKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	311.5	63.8	619	2 A97887	surface protein ps
2	311.5	63.8	619	2 A41971	surface protein ps
3	122.5	25.1	744	2 F95013	pneumococcal surf
4	115	23.6	886	2 H69378	conserved hypothet
5	110.5	22.6	896	2 S43074	epidermal growth f
6	110	22.5	1169	2 A64505	p115 homolog - Met
7	109	22.3	501	2 A44643	M protein precursor
8	108	22.1	1269	2 F84730	probable myosin he
9	104	21.3	911	2 S51441	hypothetical prote
10	102.5	21.0	897	2 A54696	Egf receptor subut
11	102.5	21.0	1190	2 E84193	chromosome segrega
12	102	20.9	2139	2 T18296	myosin heavy chain
13	101.5	20.8	1974	2 T30010	hypothetical prote
14	101	20.7	3488	2 T34418	hypothetical prote
15	100.5	20.6	764	2 T05409	hypothetical prote
16	100	20.5	1006	2 C70445	ATPase subunit of
17	99.5	20.4	1319	2 A28313	glut protein - fr
18	98.5	20.2	166	2 S73342	hypothetical prote
19	97.5	20.0	1927	2 A59236	embryonic muscle m
20	97.5	20.0	1957	2 T38077	hypothetical coile
21	97.5	20.0	1972	1 A41604	myosin heavy chain
22	97	19.9	161	2 S43396	tropomyosin TPM2 -
23	97	19.9	1093	2 S66717	hypothetical prote
24	97	19.9	1938	1 A40997	myosin heavy chain
25	96.5	19.8	522	2 G02533	occludin - human
26	96.5	19.8	629	2 T44607	hypothetical prote
27	96.5	19.8	1027	2 S37711	kinesin heavy chain
28	96	19.7	387	2 S57834	fcrA protein precu
29	96	19.7	388	2 A46173	Mrp4 protein - Str

30 96 19.7 1156 2 B70356 chromosome assembl  
31 95.5 19.6 1177 2 B75150 chromosome segrega  
32 95.5 19.6 1179 2 F71190 probable chromosom  
33 95.5 19.6 1959 1 A33977 myosin heavy chain  
34 95.5 19.6 1961 1 A61231 myosin heavy chain  
35 95.5 19.6 1999 1 S21801 myosin heavy chain  
36 95 19.5 539 2 A28549 M24 protein precur  
37 95 19.5 1138 2 T24635 hypothetical prote  
38 95 19.5 1938 2 A59293 skeletal myosin he  
39 94.5 19.4 1938 2 JC5421 smooth muscle myos  
40 94.5 19.4 1972 2 JC5420 smooth muscle myos  
41 94.5 19.4 1979 1 S03166 myosin heavy chain  
42 94 19.3 710 2 AE1956 hypothetical prote  
43 94 19.3 2116 2 A26655 myosin heavy chain  
44 94 19.3 2288 2 T29999 hypothetical prote  
45 93.5 19.2 603 2 T00379 KfAA0640 protein -

ALIGNMENTS

RESULT 1

A97887

surface protein psPa precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:9  
C:Genetics:  
A:Gene: psPa

Query Match 63.8%; Score 311.5; DB 2; Length 619;  
Best Local Similarity 65.1%; Pred. No. 7.2e-14;  
Matches 71; Conservative 8; Mismatches 11; Indels 19; Gaps 2;

QY 1 LKEIDSESDYKGERAPLOSELDKAKLSELSDKIDELDAETAKLE-----53

DB 223 LKEIDSESDYKGERAPLOSELDKAKLSELSDKIDELDAETAKLE-----282

QY .54 --KVEDFKNSDGEQGVLAAGSDLIKAKAELEKAEADLKKAVIDEPE 100

DB 283 ENNVEDYKGELEKT-----IAKAELEKTEADLKKAVIDEPE 321

RESULT 2

A41971

surface protein psPa precursor - Streptococcus pneumoniae

N:Alternate names: pneumococcal surface protein A

C:Species: Streptococcus pneumoniae

C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A>Title: Structural properties and evolutionary relationships of PsPa, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249

A:Accession: A41971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <YOT>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701

A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991  
A>Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability  
A:Reference number: A60282; MUID:91169598; PMID:2004810  
A:Accession: A60282  
A:Molecule type: protein  
A:Residues: 32-76 <TAL>  
A:Experimental source: strain JY2008  
C:Genetics:  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-619/Product: surface protein pspA #status predicted <MAT>  
F:411-430/Domain: cpl repeat homology <CP01>  
F:431-450/Domain: cpl repeat homology <CP02>  
F:451-470/Domain: cpl repeat homology <CP03>  
F:471-490/Domain: cpl repeat homology <CP04>  
F:491-510/Domain: cpl repeat homology <CP05>  
F:511-530/Domain: cpl repeat homology <CP06>  
F:531-550/Domain: cpl repeat homology <CP07>  
F:551-570/Domain: cpl repeat homology <CP08>  
F:571-591/Domain: cpl repeat homology <CP09>  
F:592-611/Domain: cpl repeat homology <CP10>  
  
Query Match 63.8%; Score 311.5; DB 2; Length 619;  
Best Local Similarity 65.1%; Pred. No. 7.2e-14;  
Matches 71; Conservative 8; Mismatches 11; Indels 19; Gaps 2;  
  
Qy 1 LKEIDESSEDYVKEGRAPLQSELDAKQAKLSKLELSKIDELDAEIAKLE----- 53  
Db 223 LKEIDESSEDYVKEGRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEQLKAAE 282  
  
Qy 54 --KVEDFNKSDGEQAGYLAAGEDLLAKKAELEKADLKKAVDPE 100  
Db 283 ENNVVEDYFKGLEKT-----IAAKKAELEKTEADLKKAVDPE 321  
  
RESULT 3  
F95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: F95013  
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: F95013  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-744 <KUR>  
A:Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:G  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0117

Query Match 25.1%; Score 122.5; DB 2; Length 744;  
Best Local Similarity 34.7%; Pred. No. 0.38;  
Matches 42; Conservative 18; Mismatches 26; Indels 35; Gaps 6;

Qy 10 EDYVKEGRAPLQSELDAKQAKLSK-----LELSDKI-----DELDAEIAKLEKDV 56  
Db 248 EAKLKKGE-----AEINAKQAELEAKQTELEKLLSLDPEGKTQDELKKEAEAELEDKKA 302  
  
Qy 57 EDFKNS-----DGEQAGYLAAGEDLLAKKAELEKADLKKAVD--EP 99  
Db 303 DELQNKVADLEKEISNLHILGGADPEDDTAALQNKLAAKKAELEKQTELEKLLSLDP 362

Qy 100 E 100  
Db 363 E 363

## RESULT 4

H69378  
conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: H69378  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Arttich, P.; Kaine, B.P.; Sykes, S.;  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: H69378  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-886 <KLE>  
A:Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB9021  
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 23.6%; Score 115; DB 2; Length 886;  
Best Local Similarity 29.1%; Pred. No. 1.4;  
Matches 39; Conservative 22; Mismatches 37; Indels 36; Gaps 4;

Qy 1 LKEIDESSEDYVKEG-----ERAPLQSELDAKQAKLSKLELSKIDELDAEIAKLEKD- 55  
Db 296 LSEINQALRDYKREGDLTREAGIQALKKAEDNSKLEITKRIEELERELERFEPKSH 355

Qy 56 --VEDFNKSDGEQAG-----QYLAAGED-----LIANKAE 84  
Db 356 RLLETLAPKMDRMOGIKAKLEKNLTPDKVKMYDLLSKAKEEKEITEKLLKLIANKSS 415

Qy 85 LEKAEADLKKAVID 98  
Db 416 LKTRGAQLKKAVID 429

## RESULT 5

S43074  
epidermal growth factor receptor substrate - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S43074; I38525  
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.  
Oncogene 9, 1039-1045, 1994  
A>Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4,  
A:Reference number: S43074; MUID:94181254; PMID:8134107  
A:Accession: S43074  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-896 <BER>  
A:Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA82305.1; PID:g4700  
R:Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner,  
Oncogene 9, 1591-1597, 1994  
A>Title: The human epr15 gene, encoding a tyrosine kinase substrate, is conserved in evo

A:Reference number: I38525; MUID:94239734; PMID:8183552  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-821, 'M', 823-896 <RES>  
A:Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260  
C:Genetics:  
A:Gene: GDB:EPS15; AF-1P; MLT5  
A:Cross-references: GDB:360337; OMIM:600051  
A:Map position: 1p32-1p32

Query Match 22.6%; Score 110.5; DB 2; Length 896;  
Best Local Similarity 27.5%; Pred. No. 2.9;  
Matches 28; Conservative 30; Mismatches 37; Indels 7; Gaps 3;  
  
Qy 3 EIDESSEDYVKE--GERAPLQSELDAKQAKLSK-----EELSKIDELDAEIAKLEKDV 56

C:Superfamily: M5 protein  
C:Keywords: coiled coil; dimer

Query Match 22.3%; Score 109; DB 2; Length 501;  
Best Local Similarity 37.0%; Pred. No. 2;  
Matches 37; Conservative 16; Mismatches 33; Indels 14; Gaps 5

QY 3 BIDSDSDYVK-EGERAPLOSELDAKQAKLSKLELSDKIDELDAETIAKLEKOVE---D 58  
Db 125 ETIEKEVEDYNKLVDKASLDKKTIESAN---SQLEFKNSQISELVAQANLNQIEKLSE 181  
QY 59 FKNSDGEAGQYLAAGE---DLIAKKAELK-----KAEAD 91  
Db 182 EKNKAEBOSQLEFKNQIADLIGCKKALEMLKAKAEBED 221

RESULT 8  
F84730  
probable myosin heavy chain [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84730  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujieuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84730  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1269 <STO>  
A:Cross-references: GB:AE002093; NID:G6598483; PIDN:AAC69932.2; GSPDB:GN00139  
C:Genetics:  
A:Gene: Atg32240  
A:Map position: 2

Query Match 22.1%; Score 108; DB 2; Length 1269;  
Best Local Similarity 34.0%; Pred. No. 6;  
Matches 35; Conservative 19; Mismatches 31; Indels 18; Gaps 4

QY 6 BSDSDYKGERAPLOSELDAKQAKLSKLE-----ELSDKIDELDAETIAKLEK 54  
Db 661 EADSKGYL--GQVAELSTLEAFQVKSSLEAALNIATENEKELTNLNAVTSBKKLEA 718  
QY 55 DVEDF--KNSDGEAGQYLAAGEDLIAKKAELKAEADIKKA 95  
Db 719 TVDYSYKISESE---NLLESIRNELNVTOGKLESIENDLKAA 758

RESULT 9  
S51441  
hypotheetical protein YLR309c - yeast. (Saccharomyces cerevisiae)  
N:Alternate names: hypotheetical protein L2142.5  
C:Species: Saccharomyces cerevisiae  
C:Date: 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C:Accession: S51441  
R:Pauley, A.  
submitted to the EMBL Data Library, November 1994  
A:Description: The sequence of S. cerevisiae cosmid L2142.  
A:Reference number: S51437  
A:Accession: S51441  
A:Molecule type: DNA  
A:Residues: 1-911 <PAU>  
A:Cross-references: UNIPROT:Q06704; EMBL:U17247; NID:G577216; PIDN:AAB67359.1;  
C:Genetics:  
A:Gene: SGD:IMH1  
A:Cross-references: SGD:S0004300; MIPS:YLR309C  
A:Map position: 12R

Query Match 21.3%; Score 104; DB 2; Length 911;  
Best Local Similarity 28.3%; Pred. No. 8;  
Matches 28; Conservative 32; Mismatches 33; Indels 6; Gaps 4

Qy 1 LKEIDF--SDSEYVKEGERAPLQSEL--DAKQAKLSKLELSDKIDELDAIAKLEKQVE 57  
Db 329 LQELQEKYKDCEDWKFQYE--DIEAEKDAKELENSQLSEKSAKELETINTLELIDTKKSLK 386  
Qy 58 DFKNSDGEQAGQYLAAGGEDLIAKKAELEKAEADLKKAV 96  
Db 387 E-KNSELSEVRDMLRTVGNELVDKDEIKSESSKQNEEV 424  
RESULT 10  
A:54696  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 09-Jul-2004  
C:Accession: A54696  
R:Fazio, F.; Minichiello, L.; Matoskova, B.; Wong, W.T.; Di Fiore, P.P.  
Mol. Cell. Biol. 13, 5814-5828, 1993  
A:Title: eps15, a novel tyrosine kinase substrate, exhibits transforming activity.  
A:Reference number: A54696; PMID:93361014; PMID:7689153  
A:Accession: A54696  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-897 <FAZ>  
A:Cross-references: UNIPROT:P42567; GB:L21768; NID:g404756; PIDN:AAA02912.1; PID:g404757  
F:48-80/Domain: calmodulin repeat homology <EF1>  
F:160-192/Domain: calmodulin repeat homology <EF2>  
F:223-255/Domain: calmodulin repeat homology <EF3>  
Query Match 21.0%; Score 102.5; DB 2; Length 897;  
Best Local Similarity 26.5%; Pred. No. 9.9;  
Matches 27; Conservative 30; Mismatches 38; Indels 7; Gaps 3;  
Qy 3 EIDESDSDYVKE--GERAPLQSELDAKQAKLSK----EELSCKIDELDAIAKLEKDV 56  
Db 353 EQDLKEKEDTVKQRTSEVQDLQDEVQRESINLKQAKQVQVQLGELDELDEKQAKLERQL 412  
Qy 57 EDFKNSDGEQAGQYLAAGGEDLIAKKAELEKAEADLKKAVDE 98  
Db 413 QEVKKKAEAE-QLTSSLKAEITTSQESQISSVEEELKAREE 453  
RESULT 11  
E84193  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: E84193  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonc  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omar, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; PMID:20504483; PMID:11016950  
A:Accession: E84193  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1190 <STO>  
A:Cross-references: UNIPROT:Q9HS95; GB:AE004437; NID:g10579965; PIDN:AAG18913.1; GSPDB:C  
C:Genetics:  
A:Gene: smc1  
Query Match 21.0%; Score 102.5; DB 2; Length 1190;  
Best Local Similarity 33.7%; Pred. No. 13;  
Matches 34; Conservative 22; Mismatches 36; Indels 9; Gaps 4;  
Qy 3 EIDESDSDYVKEGRA---PLQSELDAKQAKLSKLELSCKIDELDAIAKLEKVEDF 59  
Db 858 QADVADAEER-KADQARIEALNGDIEAKQAEAEKEA---AVEDLEAEADLKRDRER 913  
Qy 60 KN--SDGEQAGQYLAAGGEDLIAKKAELEKAEADLKKAVDE 98

Db 914 KADLSEADARDEQAAAVEDARHRLERLQRAAAQTLSSEVAE 954  
RESULT 12  
T18296  
myosin heavy chain - Entamoeba histolytica  
C:Species: Entamoeba histolytica  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18296  
R:Guillen, N.  
submitted to the EMBL Data Library, February 1997  
A:Reference number: Z18865  
A:Accession: T18296  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2139 <GUI>  
A:Cross-references: UNIPROT:Q07569; EMBL:L03534; NID:g1850912; PID:g1850913; PIDN:AA8480  
C:Genetics:  
A:Gene: mhca  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:91-780/Domain: myosin motor domain homology <MMO>  
Query Match 20.9%; Score 102; DB 2; Length 2139;  
Best Local Similarity 25.2%; Pred. No. 25;  
Matches 34; Conservative 26; Mismatches 35; Indels 40; Gaps 4;  
Qy 2 KEIDESDSDYVKEGERAPLQSELDAKQAKL-----SKLELSCKIDELDAEIA 50  
Db 1339 KSVVESKNKD--SENEKAALSEEDQANEKLNQADLRKATADLQEAENEKKAEEVAQRD 1396  
Qy 51 KLEKD-----VEDF-----KNSDGEQAGQYLAAGGEDLIAKKA 83  
Db 1397 KLVAADNKKMTKLTBEIKARDEENTYKVENYKVLKREADLEAENLDTIEKDRMNRKE 1456  
Qy 84 ELEKAEADLKKAVDE 98  
Db 1457 QVKKLEGLKETDKD 1471  
RESULT 13  
T30010  
hypothetical protein F58G4.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T30010  
R:Du, Z.; Leimbac, D.  
submitted to the EMBL Data Library, February 1996  
A:Description: The sequence of C. elegans cosmid F58G4.  
A:Reference number: Z20720  
A:Accession: T30010  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1974 <DUZ>  
A:Cross-references: UNIPROT:Q21000; EMBL:U50309; PIDN:AAB37057.1; GSPDB:GN00023; CESP:F.5  
A:Experimental source: strain Bristol N2; clone F58G4  
C:Genetics:  
A:Gene: CESP:F58G4.1  
A:Map position: 5  
A:Introns: 18/3; 111/3; 164/1; 229/1; 264/1; 378/1; 440/2; 525/3; 1177/2; 1633/3; 1863/3  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:84-776/Domain: myosin motor domain homology <MMO>  
Query Match 20.8%; Score 101.5; DB 2; Length 1974;  
Best Local Similarity 31.9%; Pred. No. 25;  
Matches 36; Conservative 18; Mismatches 24; Indels 35; Gaps 6;  
Qy 21 LQSELDAKQAKLSK-----LEELSDKIDEL-----DAEIAKLEKD 55  
Db 1128 LEELDAERNRSKAEKARNEMQMELELGDRLDEAGATQAQIELNKKREAEKLRQD 1187  
Qy 56 VEDFK-NSDGEQAG---QYLAAGE-----DLIAK-KAELEKAEADLKKAVDE 98

Db 1188 LEDAAINSETSMALRKHNDAVAELSDQLDTIQMRGLEREKNDKQREVDE 1240

REF ID: A61871  
T34418

15416

C;Species: *Caenorhabditis elegans*

**C;Accession: T34418**

submitted to the EMBL Data

all: dead; 'A  
 zed: 'A  
 zed: 'A

A;Accession: T34418

A;Molecule type: DNA

A:Cross-references: EMBL

A; Experimental source: strain Bristol N2; clone FL2F3

**A;Gene: CES**

**A; Introns: 281/3;**

Query Match 20 78: Score 101: DB 2: Length 2488.

Best Local Similarity 32.8%; Pred. No. 47;

Case	Case	Case
1	2	3
4	5	6
7	8	9
10	11	12
13	14	15
16	17	18
19	20	21
22	23	24
25	26	27
28	29	30
31	32	33
34	35	36
37	38	39
40	41	42
43	44	45
46	47	48
49	50	51
52	53	54
55	56	57
58	59	60
61	62	63
64	65	66
67	68	69
70	71	72
73	74	75
76	77	78
79	80	81
82	83	84
85	86	87
88	89	90
91	92	93
94	95	96
97	98	99
100	101	102
103	104	105
106	107	108
109	110	111
112	113	114
115	116	117
118	119	120
121	122	123
124	125	126
127	128	129
130	131	132
133	134	135
136	137	138
139	140	141
142	143	144
145	146	147
148	149	150
151	152	153
154	155	156
157	158	159
160	161	162
163	164	165
166	167	168
169	170	171
172	173	174
175	176	177
178	179	180
181	182	183
184	185	186
187	188	189
190	191	192
193	194	195
196	197	198
199	200	201
202	203	204
205	206	207
208	209	210
211	212	213
214	215	216
217	218	219
220	221	222
223	224	225
226	227	228
229	230	231
232	233	234
235	236	237
238	239	240
241	242	243
244	245	246
247	248	249
250	251	252
253	254	255
256	257	258
259	260	261
262	263	264
265	266	267
268	269	270
271	272	273
274	275	276
277	278	279
280	281	282
283	284	285
286	287	288
289	290	291
292	293	294
295	296	297
298	299	300
301	302	303
304	305	306
307	308	309
310	311	312
313	314	315
316	317	318
319	320	321
322	323	324
325	326	327
328	329	330
331	332	333
334	335	336
337	338	339
340	341	342
343	344	345
346	347	348
349	350	351
352	353	354
355	356	357
358	359	360
361	362	363
364	365	366
367		

DD 1009 NEIDENKLDAAETAAKINQEADEKSKLDA-QEKIKRVSEDDAARKEKELNDKT-KLESEL 1060

QY 50 A-----KLEKDVEDFKNSDGEQAGQY-----LAAAGEDLIAKKAEELEKAEA 90

Db 1067 ATKASADKLKLEEQAKKAAEVEAAKKQKEKDEQLKLDTEAASKKAAAEKLELEK-QA 1125

91 DLKKA 95  
Ov

1126 01442 1130  
1126 01442 1130

## RESULT 15

hypothetical protein F10M6.170 - *Arabidopsis thaliana*

C/Species: ADAPTODONTIS CHALIANA (mouse-eat cress)  
C/Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #cont changes 00 Jul 2001

C;Accession: T05409

submitted to the Protein Sequence Database, February 1998

A:Accession: T05409  
 A:Accession: T05411

A; molecule type: DNA

A;Cross-references: UNIPROT:O49371; EMBL:AL021811

C;Genetics:

A:Map Position: 4  
A:Note: F10M6 170

Best Local Similarity 30.7%; Pred. No. 11;

1  
2  
2  
3  
3  
4  
4  
5  
5  
6  
6  
7  
7  
8  
8  
9  
9  
10  
10  
11  
11  
12  
12  
13  
13  
14  
14  
15  
15  
16  
16  
17  
17  
18  
18  
19  
19  
20  
20  
21  
21  
22  
22  
23  
23  
24  
24  
25  
25  
26  
26  
27  
27  
28  
28  
29  
29  
30  
30  
31  
31  
32  
32  
33  
33  
34  
34  
35  
35  
36  
36  
37  
37  
38  
38  
39  
39  
40  
40  
41  
41  
42  
42  
43  
43  
44  
44  
45  
45  
46  
46  
47  
47  
48  
48  
49  
49  
50  
50  
51  
51  
52  
52  
53  
53  
54  
54  
55  
55  
56  
56  
57  
57  
58  
58  
59  
59  
60  
60  
61  
61  
62  
62  
63  
63  
64  
64  
65  
65  
66  
66  
67  
67  
68  
68  
69  
69  
70  
70  
71  
71  
72  
72  
73  
73  
74  
74  
75  
75  
76  
76  
77  
77  
78  
78  
79  
79  
80  
80  
81  
81  
82  
82  
83  
83  
84  
84  
85  
85  
86  
86  
87  
87  
88  
88  
89  
89  
90  
90  
91  
91  
92  
92  
93  
93  
94  
94  
95  
95  
96  
96  
97  
97  
98  
98  
99  
99  
100  
100  
101  
101  
102  
102  
103  
103  
104  
104  
105  
105  
106  
106  
107  
107  
108  
108  
109  
109  
110  
110  
111  
111  
112  
112  
113  
113  
114  
114  
115  
115  
116  
116  
117  
117  
118  
118  
119  
119  
120  
120  
121  
121  
122  
122  
123  
123  
124  
124  
125  
125  
126  
126  
127  
127  
128  
128  
129  
129  
130  
130  
131  
131  
132  
132  
133  
133  
134  
134  
135  
135  
136  
136  
137  
137  
138  
138  
139  
139  
140  
140  
141  
141  
142  
142  
143  
143  
144  
144  
145  
145  
146  
146  
147  
147  
148  
148  
149  
149  
150  
150  
151  
151  
152  
152  
153  
153  
154  
154  
155  
155  
156  
156  
157  
157  
158  
158  
159  
159  
160  
160  
161  
161  
162  
162  
163  
163  
164  
164  
165  
165  
166  
166  
167  
167  
168  
168  
169  
169  
170  
170  
171  
171  
172  
172  
173  
173  
174  
174  
175  
175  
176  
176  
177  
177  
178  
178  
179  
179  
180  
180  
181  
181  
182  
182  
183  
183  
184  
184  
185  
185  
186  
186  
187  
187  
188  
188  
189  
189  
190  
190  
191  
191  
192  
192  
193  
193  
194  
194  
195  
195  
196  
196  
197  
197  
198  
198  
199  
199  
200  
200  
201  
201  
202  
202  
203  
203  
204  
204  
205  
205  
206  
206  
207  
207  
208  
208  
209  
209  
210  
210  
211  
211  
212  
212  
213  
213  
214  
214  
215  
215  
216  
216  
217  
217  
218  
218  
219  
219  
220  
220  
221  
221  
222  
222  
223  
223  
224  
224  
225  
225  
226  
226  
227  
227  
228  
228  
229  
229  
230  
230  
231  
231  
232  
232  
233  
233  
234  
234  
235  
235  
236  
236  
237  
237  
238  
238  
239  
239  
240  
240  
241  
241  
242  
242  
243  
243  
244  
244  
245  
245  
246  
246  
247  
247  
248  
248  
249  
249  
250  
250  
251  
251  
252  
252  
253  
253  
254  
254  
255  
255  
256  
256  
257  
257  
258  
258  
259  
259  
260  
260  
261  
261  
262  
262  
263  
263  
264  
264  
265  
265  
266  
266  
267  
267  
268  
268  
269  
269  
270  
270  
271  
271  
272  
272  
273  
273  
274  
274  
275  
275  
276  
276  
277  
277  
278  
278  
279  
279  
280  
280  
281  
281  
282  
282  
283  
283  
284  
284  
285  
285  
286  
286  
287  
287  
288  
288  
289  
289  
290  
290  
291  
291  
292  
292  
293  
293  
294  
294  
295  
295  
296  
296  
297  
297  
298  
298  
299  
299  
300  
300  
301  
301  
302  
302  
303  
303  
304  
304  
305  
305  
306  
306  
307  
307  
308  
308  
309  
309  
310  
310  
311  
311  
312  
312  
313  
313  
314  
314  
315  
315  
316  
316  
317  
317  
318  
318  
319  
319  
320  
320  
321  
321  
322  
322  
323  
323  
324  
324  
325  
325  
326  
326  
327  
327  
328  
328  
329  
329  
330  
330  
331  
331  
332  
332  
333  
333  
334  
334  
335  
335  
336  
336  
337  
337  
338  
338  
339  
339  
340  
340  
341  
341  
342  
342  
343  
343  
344  
344  
345  
345  
346  
346  
347  
347  
348  
348  
349  
349  
350  
350  
351  
351  
352  
352  
353  
353  
354  
354  
355  
355  
356  
356  
357  
357  
358  
358  
359  
359  
360  
360  
361  
361  
362  
362  
363  
363  
364  
364  
365  
365  
366  
366  
367  
367  
368  
368  
369  
369  
370  
370  
371  
371  
372  
372  
373  
373  
374  
374  
375  
375  
376  
376  
377  
377  
378  
378  
379  
379  
380  
380  
381  
381  
382  
382  
383  
383  
384  
384  
385  
385  
386  
386  
387  
387  
388  
388  
389  
389  
390  
390  
391  
391  
392  
392  
393  
393  
394  
394  
395  
395  
396  
396  
397  
397  
398  
398  
399  
399  
400  
400  
401  
401  
402  
402  
403  
403  
404  
404  
405  
405  
406  
406  
407  
407  
408  
408  
409  
409  
410  
410  
411  
411  
412  
412  
413  
413  
414  
414  
415  
415  
416  
416  
417  
417  
418  
418  
419  
419  
420  
420  
421  
421  
422  
422  
423  
423  
424  
424  
425  
425  
426  
426  
427  
427  
428  
428  
429  
429  
430  
430  
431  
431  
432

QY 2 NEIDESUSEDIVNEGERAPLQSELDJAKQANLSK-----EELSUKIDELJAEIAKLE 53

Db 163 REIEELKHKLREDEREAAALQSSSLTLKEEELEKMRQEIANSRKEVSMASIEFESKSQLS 222

QY 54 KDVDFKNSDGE--QAGQYLAAGEDLIAKKA-----ELEKAEADLKKAVDE 98

223 KANEVSKROEGETYALORALEEKEEELIETSATKKEEET.BETEAN.KKOTEE 276

Search completed: June 18, 2005, 17:03:52

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-4  
Perfect score: 488  
Sequence: 1 LKEIDSDSDYVKEGERAP.....KKALEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03:.\*  
1: uniprot\_sprot;.\*  
2: uniprot\_trembl;.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	96.7	394	Q9LAY6	Q9lay6 streptococc
2	472	96.7	395	Q9LAZ1	Q9laz1 streptococc
3	459	94.1	406	Q9LAZ0	Q9laz0 streptococc
4	452	92.6	340	Q8KQK5	Q8kqk5 streptococc
5	438	89.8	416	Q9LAY8	Q9lay8 streptococc
6	435	89.1	225	Q9L591	Q9l591 streptococc
7	431	88.3	194	Q9L5B5	Q9l5b5 streptococc
8	431	88.3	218	Q6UEB2	Q6ueb2 streptococc
9	431	88.3	233	Q9L568	Q9l568 streptococc
10	431	88.3	236	Q9L569	Q9l569 streptococc
11	431	88.3	243	Q9L564	Q9l564 streptococc
12	431	88.3	243	Q9L567	Q9l567 streptococc
13	431	88.3	244	Q9L565	Q9l565 streptococc
14	431	88.3	247	Q9L566	Q9l566 streptococc
15	431	88.3	249	Q9L570	Q9l570 streptococc
16	431	88.3	254	Q9L563	Q9l563 streptococc
17	431	88.3	401	Q9LAZ2	Q9laz2 streptococc
18	430	88.1	222	Q9L577	Q9l577 streptococc
19	430	88.1	262	Q9L576	Q9l576 streptococc
20	430	88.1	415	Q9LAY7	Q9lay7 streptococc
21	423	86.7	255	Q9L581	Q9l581 streptococc
22	423	86.7	255	Q9L5B6	Q9l5b6 streptococc
23	420	86.1	246	Q9L578	Q9l578 streptococc
24	403	82.6	393	Q9LAZ3	Q9laz3 streptococc
25	392	80.3	237	Q8GNS9	Q8gns9 streptococc
26	386	79.1	237	Q9L592	Q9l592 streptococc
27	386	79.1	395	Q9LAY9	Q9lay9 streptococc
28	327.5	67.1	417	Q9LAY3	Q9lay3 streptococc
29	325.5	66.7	739	Q9RQT4	Q9rqt4 streptococc
30	325.5	66.7	820	Q9RQT1	Q9rqt1 streptococc
31	325.5	66.7	929	Q9KK19	Q9kk19 streptococc

32 325.5 66.7 929 2 Q9ZAY5 streptococc  
33 311.5 63.8 619 2 Q54972 streptococc  
34 311.5 63.8 619 2 Q8DR10 streptococc  
35 310.5 63.6 249 2 Q9L575 streptococc  
36 310.5 63.6 415 2 Q9LAY1 streptococc  
37 309.5 63.4 437 2 Q9LAY4 streptococc  
38 306.5 62.8 99 2 Q8QKQ4 streptococc  
39 303.5 62.2 224 2 Q8GNS8 streptococc  
40 303.5 62.2 426 2 Q9LAY5 streptococc  
41 298.5 61.2 395 2 Q9LAY2 streptococc  
42 298.5 61.2 408 2 Q9LAY0 streptococc  
43 290 59.4 869 2 Q9KK27 streptococc  
44 173.5 35.6 481 2 Q9LAX5 streptococc  
45 173 35.5 479 2 Q9LAX2 streptococc

## ALIGNMENTS

## RESULT 1

Q9LAY6 PRELIMINARY; PRT; 394 AA.  
ID AC Q9LAY6;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L81905;  
RX MEDLINE=20448953; PubMed=10992499;  
RX DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071809; AAF27705.1; -;  
DR InterPro; IPR009082; His\_kin\_homodim.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 394 394  
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C88FAA CRC64;

Query Match 96.7%; Score 472; DB 2; Length 394;  
Best Local Similarity 98.0%; Pred. No. 2.1e-22;  
Matches 98; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPQSELDQAQAKLSLELSDKIDELDAETAKLEKVEDFK 60  
DB 213 LKEIDSDSDYVKEGERAPQSELDQAQAKLSLELSDKIDELDAETAKLEKVEDFK 272

QY 61 NSDGEQAGQYLAAGEDIKKALEKAEADLKKAVIDEPE 100  
DB 273 NSDGEQAGQYLAAGEDIKKALEKAEADLKKAVIDEPE 312

## RESULT 2

Q9LAZ1 PRELIMINARY; PRT; 395 AA.  
ID AC Q9LAZ1;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```
OC Streptococcus
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9739;
RX MEDLINE=20448953; PubMed=10992499;
EX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27700.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECAC41DB7F95 CRC64;

Query Match 96.7%; Score 472; DB 2; Length 395;
Best Local Similarity 98.0%; Pred. No. 2.1e-22;
Matches 98; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 213 LKEIDESDSEYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 272

Qy 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 100
Db 273 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 312

RESULT 3
Q9LAZO PRELIMINARY; PRT; 406 AA.
AC Q9LAZO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DL6A;
RX MEDLINE=20448953; PubMed=10992499;
EX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 94.1%; Score 459; DB 2; Length 406;
Best Local Similarity 95.0%; Pred. No. 1.4e-21;
Matches 95; Conservative 2; Mismatches 23; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 213 LKEIDESDSEYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 272

Qy 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 100
Db 273 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 312
```

```
RESULT 4
Q8KQK5 PRELIMINARY; PRT; 340 AA.
AC Q8KQK5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
EX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
immunization with DNA vaccines against Streptococcus pneumoniae
expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1
FT NON TER 340
SQ SEQUENCE 340 AA; 38023 MW; EE07ECF00B1PBD57 CRC64;

Query Match 92.6%; Score 452; DB 2; Length 340;
Best Local Similarity 93.0%; Pred. No. 3.4e-21;
Matches 93; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Db 197 LKEIDESDSEYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 256

Qy 61 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 100
Db 257 NSDGEQAGQYLAAGEDIKAKAELEKAEADLKKAADVEPE 296

RESULT 5
Q9LAY8 PRELIMINARY; PRT; 416 AA.
AC Q9LAY8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8838;
RX MEDLINE=20448953; PubMed=10992499;
EX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071807; AAF27703.1; -
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 416
FT NON TER 416
SQ SEQUENCE 416 AA; 45987 MW; 990C8858BC6B12C7 CRC64;
```





```
RESULT 9
Q9L568
ID Q9L568 PRELIMINARY; PRT; 233 AA.
AC Q9L568;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255902; AAF70092.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 233
SQ SEQUENCE 233 AA; 24514 MW; D5C494019C45BFE2 CRC64;

Query Match 88.3%; Score 431; DB 2; Length 233;
Best Local Similarity 89.0%; Pred. No. 5.2e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 87

Qy 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 88 NSDGEQAEQYLVAAKDKLDAKKAELTEADLKKAVIDEPE 127

RESULT 10
Q9L569
ID Q9L569 PRELIMINARY; PRT; 236 AA.
AC Q9L569;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=177;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=177;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70096.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT InterPro; IPR00533; Tropomyosin.
FT PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 88.3%; Score 431; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 5.4e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 133

Qy 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 134 NSDGEQAEQYLVAAKDKLDAKKAELTEADLKKAVIDEPE 173

RESULT 12
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
```

```
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255901; AAF70091.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 236
SQ SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;

Query Match 88.3%; Score 431; DB 2; Length 236;
Best Local Similarity 89.0%; Pred. No. 5.2e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 49 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 108

Qy 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 109 NSDGEQAEQYLVAAKDKLDAKKAELTEADLKKAVIDEPE 148

RESULT 11
Q9L564
ID Q9L564 PRELIMINARY; PRT; 243 AA.
AC Q9L564;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70096.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT InterPro; IPR00533; Tropomyosin.
FT PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 88.3%; Score 431; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 5.4e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 LKEIDESDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 133

Qy 61 NSDGEQAGQYLAAGEDIKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 134 NSDGEQAEQYLVAAKDKLDAKKAELTEADLKKAVIDEPE 173

RESULT 12
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255903; AAF70093.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match 88.3%; Score 431; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 5.4e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKQVDEFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 50 LKEIDESDSYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKQVDEFK 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 NSDGEQAEQYLVAAKDKLDKAKAELENTADLKKAVIDEPE 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
QY1565 PRELIMINARY; PRT; 244 AA.
AC Q9L565;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=183;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=183;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255905; AAF70095.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 1

Query Match 88.3%; Score 431; DB 2; Length 247;
Best Local Similarity 89.0%; Pred. No. 5.4e-20;
Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKQVDEFK 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 72 LKEIDESDSYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKQVDEFK 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NSDGEQAGQYLAAGEDLIAKAELEKAEADLKKAVIDEPE 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 132 NSDGEQAEQYLVAAKDKLDKAKAELENTADLKKAVIDEPE 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
QY1570 PRELIMINARY; PRT; 249 AA.
AC Q9L570;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
FT NON_TER 1
```

OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=137;  
 EX MEDLINE=20472698; PubMed=11015380;  
 RA Beall B., Gherardi G., Facklam R.K., Hollingshead S.K.;  
 RT "Pneumococcal psa sequence types of prevalent multiresistant  
 pneumococcal strains in the United States and of internationally  
 disseminated clones";  
 RL J. Clin. Microbiol. 38:3663-3669(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=137;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF255900; AAF70090.1; -  
 DR InterPro; IPR009082; His\_kin\_homodim.  
 DR InterPro; IPR005533; Tropomyosin.  
 DR PRINTS; PR00194; TROPOMYOSIN.  
 FT NON\_TER 1  
 FT NON\_TER 249  
 SQ SEQUENCE 249 AA; 26262 MW; E365548931381011 CRC64;

Query Match 88.3%; Score 431; DB 2; Length 249;  
 Best Local Similarity 89.0%; Pred. No. 5.5e-20;  
 Matches 89; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 LKEIDESDSEYVKEGERAPLOSLEDAKQAKLSKLEELSDKIDELDAEIAKLEKXVEDFK 60  
 Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 50 LKEIDESDSEYIKEGLRAPLQSLDAKKAQKLSKLEELSDKIDELDAEIAKLEKXVEDFK 109  
 QY 61 NSDGEAQGLAAAGEDIKAKAELEKAEADLKXAVDEPE 100  
 Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 110 NSDGEAQGLVAKKDLDAKKAELENTADLKXAVDEPE 149

Search completed: June 18, 2005, 17:01:34  
 Job time : 60.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-5  
Perfect score: 490  
Sequence: 1 LKXIDSDSDYKGERAP.....KXALEQTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	94.5	194	2 AAW14584	AAW14584 Streptococ
2	463	94.5	194	7 ABW02618	Abw02618 Db16ac pn
3	463	94.5	8991	6 ABU08487	Abu08487 S. pneumo
4	450	91.8	183	2 AAW14570	AAW14570 Streptoco
5	450	91.8	183	7 ABW02604	Abw02604 Bg9739c p
6	448	91.4	168	7 ABW02609	Abw02609 Lb1905c p
7	438	89.4	550	8 ADK48356	Adk48356 Streptoco
8	438	89.4	550	8 ADR95223	Adr95223 Novel S.
9	432.5	88.3	167	2 AAW14575	AAW14575 Streptoco
10	410	83.7	166	2 AAW14568	AAW14568 Streptoco
11	410	83.7	166	7 ABW02602	Abw02602 Bg8743c p
12	392.5	80.1	185	7 AAW14566	AAW14566 Streptoco
13	392.5	80.1	185	7 ABW02600	Abw02600 Ac94c pne
14	334.5	68.3	204	2 AAW14571	AAW14571 Streptoco
15	334.5	68.3	204	7 ABW02605	Abw02605 Bf1019c p
16	318.5	65.0	198	7 ABW02615	Abw02615 Rxlcl pneu
17	318.5	65.0	315	2 AAY04375	Aay04375 Streptoco
18	318.5	65.0	619	2 AAR63437	Aar63437 Pneumococ
19	318.5	65.0	619	2 AAR87598	Aar87598 Pneumococ
20	318.5	65.0	619	2 AAR86911	Aar86911 Pneumococ
21	318.5	65.0	619	2 AAY14838	Aay14838 Streptoco
22	318.5	65.0	619	5 AAE18782	AAE18782 S. pneumo
23	318.5	65.0	619	6 ABW45778	Abw45778 protein e
24	318.5	65.0	619	8 AD052126	Ado52126 Streptoco
25	318.5	65.0	648	2 AAW70336	Aaw70336 Pneumococ

26	318.5	65.0	648	2 AAW62274	AAW62274 Streptoco
27	318.5	65.0	648	2 AAY1837	Aay1837 Streptoco
28	318.5	65.0	648	2 AAW87879	AAW87879 A. pneumoc
29	318.5	65.0	653	2 AAW92456	AAW92456 S. pneumo
30	318.5	65.0	684	2 AAR73912	Aar73912 Streptoco
31	317.5	64.8	198	2 AAW14581	AAW14581 Streptoco
32	314.5	64.2	170	7 ABW02614	Abw02614 Rct135c p
33	314.5	64.2	181	7 ABW02596	Abw02596 0922134c
34	314.5	64.2	865	6 ABU08489	Abu08489 S. pneumo
35	314.5	64.2	929	2 AAW14593	AAW14593 Streptoco
36	314.5	64.2	929	2 AAY43384	Aay43384 S. pneumo
37	311.5	63.6	188	2 AAW14580	AAW14580 Streptoco
38	311.5	63.6	188	7 ABW02613	Abw02613 Rct129c p
39	309.5	63.2	195	2 AAW14591	AAW14591 Streptoco
40	309.5	63.2	195	7 ABW02625	Abw02625 Wu2c pneu
41	302.5	61.7	588	6 ABU08491	Abu08491 Coiled co
42	302.5	61.7	589	2 AAY43392	Aay43392 PspC alph
43	300.5	61.3	204	2 AAW14578	AAW14578 Streptoco
44	300.5	61.3	204	7 ABW02612	Abw02612 Rct123c p
45	300	61.2	180	2 AAW14562	AAW14562 Streptoco

## ALIGNMENTS

RESULT 1  
AAW14584  
ID AAW14584 standard; protein; 194 AA.  
XX AC AAW14584;  
XX  
DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
DE Streptococcus pneumoniae PspA central region.  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
OS Streptococcus pneumoniae; strain Db16.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 61 /note= "unidentified amino acid"  
XX  
XX WO9709994-Al.  
XX  
XX 20-MAR-1997.  
XX  
XX 16-SEP-1996; 96WO-US014819.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
XX Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.  
XX  
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
XX in vaccines for protecting animals against S.pneumoniae infection.  
XX  
XX Example 6; Fig 13; 296pp; English.

This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain Db16. Comparison of the N-terminal and central regions (AAW14533-57 and CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can be used to divide the strains into several families based on sequence homologies. PspA polypeptides, or fragments of them, can be used in vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX  
SQ Sequence 194 AA;

Query Match 94.5%; Score 463; DB 2; Length 194;  
Best Local Similarity 97.0%; Pred. No. 8.5e-34;  
Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 LKEIDSDSDYVYKGERAPLQSELDKAKLQKLSKLELSKDIDELDAEIAKLEKDVDFK 60  
Db 1 LKEIDSDSDYVYKGERAPLQSELDKAKLQKLSKLELSKDIDELDAEIAKLEKDVDFK 60  
Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 99  
Db 61 XSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 99

RESULT 2  
ABW02618  
ID ABW02618 standard; protein; 194 AA.  
XX  
AC ABW02618;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Db16ac pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.

Key Location/Qualifiers  
FH Misc-difference 1..194  
FT /note= "Xaa = Unknown amino acid"  
FT  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
XX WPI; 2003-862841/80.  
XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 64; 121pp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal  
CC surface protein A (PspA) central region. This sequence is used in the  
CC exemplification of the invention  
XX  
SQ Sequence 194 AA;

Query Match 94.5%; Score 463; DB 7; Length 194;  
Best Local Similarity 97.0%; Pred. No. 8.5e-34;  
Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 LKEIDSDSDYVYKGERAPLQSELDKAKLQKLSKLELSKDIDELDAEIAKLEKDVDFK 60  
Db 1 LKEIDSDSDYVYKGERAPLQSELDKAKLQKLSKLELSKDIDELDAEIAKLEKDVDFK 60  
Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 99  
Db 61 XSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 99

RESULT 3  
ABU08487  
ID ABU08487 standard; protein; 8991 AA.  
XX  
AC ABU08487;  
XX  
DT 24-JUN-2003 (first entry)  
XX  
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
XX  
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
KW antibacterial.

XX Streptococcus pneumoniae.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 1..8991  
FT /note= "All Xaa residues within this sequence are  
FT unknown"

XX  
XX US6500613-B1.  
XX  
PD 31-DEC-2002.  
XX  
PF 16-SEP-1996; 96US-00714741.  
XX  
PR 15-SEP-1995; 95US-00529055.  
XX  
PA (UYAL-) UNIV ALABAMA.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
XX Hollingshead S, Tart R, Brooks-Walter A;  
XX  
XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
PT hybridization probe, comprises pneumococcal surface protein having alpha-  
PT helical, proline rich and repeat regions.  
XX  
XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus  
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA) -  
CC like protein having alpha-helical, proline rich and repeat regions. The  
CC PspC and PspA proteins may be used in a vaccine to protect against  
CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
CC PspA may be used for the expression of the proteins, and as PCR primers  
CC or hybridisation probes. The present sequence represents S. pneumoniae  
XX PspA protein  
XX Sequence 8991 AA;

Query Match 94.5%; Score 463; DB 6; Length 8991;  
Best Local Similarity 97.0%; Pred. No. 6.8e-32;  
Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60  
DB 7537 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 7596

QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEP 99  
DB 7597 XSDGEQAGQYLAABEDLIAKAELEQTEADLKAVNEP 7635

RESULT 4  
AAW14570  
ID AAW14570 standard; protein; 183 AA.  
XX AC AAW14570;  
DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX STREPTOCOCCUS PNEUMONIAE PspA central region.  
DE DE  
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
OS STREPTOCOCCUS PNEUMONIAE; strain Bg9739.  
XX W09709994-A1.  
XX PD 20-MAR-1997.  
XX PF 16-SEP-1996; 96WO-US014819.  
XX PR 15-SEP-1995; 95US-00529055.  
XX PA (UABR-) UAB RES FOUND.  
XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.

XX STREPTOCOCCUS PNEUMONIAE surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX Example 6; Fig 13; 296pp; English.  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX Sequence 183 AA;

Query Match 91.8%; Score 450; DB 2; Length 183;  
Best Local Similarity 93.0%; Pred. No. 1.2e-32;  
Matches 93; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
SQ Sequence 183 AA;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60  
DB 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60

QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEP 100  
DB 61 NSDGEQAGQYLAABEDLIAKAELEKAEADLKAVDEP 100

RESULT 5  
ABW02604  
ID ABW02604 standard; protein; 183 AA.  
XX AC ABW02604;  
DT 12-FEB-2004 (first entry)  
XX Bg9739c pneumococcal surface protein A (PspA) central region.  
DE DE  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX Unidentified.  
XX OS US6592876-B1.  
XX PN 15-JUL-2003.  
XX PD 15-SEP-1995; 95US-00529055.  
XX PF 20-APR-1993; 93US-00048896.  
XX PR 06-JUN-1995; 95US-00465746.  
XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX Example 6; SEQ ID NO 50; 121pp; English.

XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Bg9739c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention

XX Query Match 91.8%; Score 450; DB 7; Length 183;  
Best Local Similarity 93.0%; Pred. No. 1.2e-32;  
Matches 93; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
SQ Sequence 183 AA;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60  
DB 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60  
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEP 100  
DB 61 NSDGEQAGQYLAABEDLIAKAELEKAEADLKAVDEP 100

RESULT 6

```
ABW02609
ID ABW02609 standard; protein; 168 AA.
XX
AC ABW02609;
XX
DT 12-FEB-2004 (first entry)
XX
DE L81905c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..168
FT /note= "Xaa = Unknown amino acid"
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 55; 131pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is L81905c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX
XX Sequence 168 AA;
XX
XX Query Match 91.4%; Score 448; DB 7; Length 168;
XX Best Local Similarity 93.0%; Pred. No. 1.6e-32;
XX Matches 93; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 LKEDISSDYVKEGERAPLOSELDAKQAKLSKLELSDKIDELDAIAKLEKXVEDFK 60
XX
XX 1 LKEDISSDYVKEGERAPLOSELDAKQAKLSKLELSXSDKXDELDAIAKLEKXVEDFK 60
XX
XX 61 NSDGEQAGQYLAAAEEDLIAKAELEQTEADLKAVHEPE 100
XX
XX 61 NSDGEQAGQYLAAAEEDLIAKAELEQTEADLKAVDEPE 100
XX
XX
XX RESULT 7
XX ADK48356
XX ID ADK48356 standard; protein; 550 AA.
XX
```

---

```
XX
AC ADK48356;
XX
DT 20-MAY-2004 (first entry)
XX
DE Streptococcus pneumoniae protein, Seq ID No 4871.
XX
KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX
OS Streptococcus pneumoniae.
XX
PN US6699703-B1.
XX
PD 02-MAR-2004.
XX
PF 26-MAY-2000; 2000US-00583110.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX WPI; 2004-212399/20.
XX
XX N-PSDB; ADK45695.
XX
XX New nucleic acid molecules and polypeptides useful for diagnosing,
XX preventing and treating pathological conditions resulting from bacterial
XX infection, e.g. Streptococcus pneumoniae infection, and in drug
XX screening.
XX
XX Disclosure; SEQ ID NO 4871; 301pp; English.
XX
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
XX and polypeptides. The nucleic acids and proteins are useful for
XX diagnosing, preventing and treating pathological conditions resulting
XX from bacterial infection, such as S. pneumoniae infection. These may also
XX be used for drug screening procedures. The present sequence represents a
XX Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
XX data for this patent did not appear in the printed specification but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html.
XX
XX Sequence 550 AA;
XX
XX Query Match 89.4%; Score 438; DB 8; Length 550;
XX Best Local Similarity 91.0%; Pred. No. 4.9e-31;
XX Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 LKEDISSDYVKEGERAPLOSELDAKQAKLSKLELSDKIDELDAIAKLEKXVEDFK 60
XX
XX 144 LKEDISSDYVKEGERAPLOSELDAKQAKLSKLELSDKIDELDAIAKLEKXVEDFK 203
XX
XX 61 NSDGEQAGQYLAAAEEDLIAKAELEQTEADLKAVHEPE 100
XX
XX 204 NSNGEQAEQYRAAEEDLIAKQAELEKTEADLKAVNEPE 243
XX
XX
XX RESULT 8
XX ID ADK95223 standard; protein; 550 AA.
XX
XX AC ADK95223;
XX
XX 16-DEC-2004 (first entry)
XX
XX
XX Novel S. pneumoniae protein sequence, SEQ ID 3858.
XX
XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
XX bacterial infection.
XX
```



OS Streptococcus pneumoniae.

XX US6800744-B1.

XX 05-OCT-2004.

XX 30-JUN-1998; 98US-00107433.

XX 02-JUL-1997; 97US-0051553P.

PR 12-MAY-1998; 98US-0085131P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2004-697205/68.

DR N-PSDB; ADR92620.

XX New isolated nucleic acid encoding a Streptococcus pneumoniae polypeptide, useful for diagnosing, preventing and/or treating pathological conditions resulting from the bacterial infection.

XX Disclosure; SEQ ID NO 3858; 151pp; English.

XX The invention relates to an isolated nucleic acid comprising a sequence encoding a Streptococcus pneumoniae ADR91366polypeptide, or its fragments, with any of 9 fully defined sequences (appearing as ADR94308, ADR94489, ADR94800, ADR94969, ADR95253, ADR95642, ADR95682, ADR96079) or any of the fully defined sequences appearing as ADR91705, ADR91886, ADR92197, ADR92234, ADR93039, ADR92366, ADR92650 or ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide sequences, or at least 40, 60 or 300 consecutive nucleotides, which is hybridisable under high stringency conditions to the nucleotide sequence. The nucleic acids and proteins are chosen from 5206 disclosed sequences. CC Also included are a recombinant expression vector comprising the isolated nucleic acid cited above operably linked to a transcription regulatory element, a cell comprising the recombinant expression vector and a probe comprising at least 20 consecutive nucleotides of the nucleotide sequences as cited above. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of pathological conditions resulting from bacterial infection by Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and otitis media. The present sequence is one of the 2603 disclosed S. pneumoniae protein sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6800744B1.

XX Sequence 550 AA;

Query Match 89.4%; Score 438; DB 8; Length 550;  
Best Local Similarity 91.0%; Pred. No. 4.9e-31;  
Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60

Db 144 LKEIDSESDYVKEGLRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 203

QY 61 NSDGEQAGQYLAAREEDLIKKAELQTEADLKKAVHEPE 100

Db 204 NSNGEQAGQYLAAREEDLIKKAELQTEADLKKAVHEPE 243

RESULT 9

AAW14575

ID AAW14575 standard; protein; 167 AA.

XX AAW14575;

XX AAW14575;

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

DE Streptococcus pneumoniae PspA central region.

XX

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis; bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain L81905.

XX Key Location/Qualifiers

FT Misc-difference 37

FT /note= "unidentified amino acid"

FT Misc-difference 41

FT /note= "unidentified amino acid"

FT Misc-difference 83

FT /note= "unidentified amino acid"

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used

PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the alpha-helix region and some of the proline-rich region, of pneumococcal surface protein A (PspA) of Streptococcus pneumoniae strain AAL81905. CC Comparison of the N-terminal and central regions (AAW14533-57 and CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can CC be used to divide the strains into several families based on sequence CC homologies. PspA polypeptides, or fragments of them, can be used in CC vaccines to protect animals against S. pneumoniae infection and hence for CC the prevention of diseases such as otitis media, meningitis, bacteraemia CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical CC region and the immediate 5' tip of the coding sequence are likely to be CC the critical sequences for predicting PspA cross-reactions and vaccine CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 167 AA;

Query Match 88.3%; Score 432.5; DB 2; Length 167;  
Best Local Similarity 92.0%; Pred. No. 3.9e-31;  
Matches 92; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEKVEDFK 60

Db 1 LKEIDSESDYVKEGFRAPLQSELDKQAKLS-LEESDKQDELDAETAKLEKVEDFK 59

QY 61 NSDGEQAGQYLAAREEDLIKKAELQTEADLKKAVHEPE 100

Db 60 NSDGEQAGQYLAAREEDLIKKAELQTEADLKKAVHEPE 99

RESULT 10

AAW14568

ID AAW14568 standard; protein; 166 AA.

XX AAW14568;

XX AAW14568;

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.



PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX SQ Sequence 185 AA;  
 Query Match 80.1%; Score 392.5; DB 2; Length 185;  
 Best Local Similarity 82.2%; Pred. No. 1.7e-27;  
 Matches 83; Conservative 7; Mismatches 10; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLELSKDIDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDESDSYVKEGLRVPLOSELDAKQAKLSKLELSKDIDELDAETAK-LEKDVDF 60  
 QY 60 KNSDGEQAGYLAAREEDLIAKAELEQTEADLKKAHVEPE 100  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELEKTEADLKKAHVEPE 101  
 RESULT 13  
 ABW02600  
 ID ABW02600 standard; protein; 185 AA.  
 AC ABW02600;  
 XX 12-FEB-2004 (first entry)  
 XX Ac94c pneumococcal surface protein A (PspA) central region.  
 DE Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 OS US6592876-B1.  
 XX US6592876-B1.  
 XX 15-JUL-2003.  
 XX 15-SEP-1995; 95US-00529055.  
 XX 20-APR-1993; 93US-00048896.  
 XX 06-JUN-1995; 95US-00465746.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain.

PT comprises at least two different full length isolated gene encoding  
 XX pneumococcal surface protein A.  
 PS Example 6; SEQ ID NO 46; 121pp; English.  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX SQ Sequence 185 AA;  
 Query Match 80.1%; Score 392.5; DB 7; Length 185;  
 Best Local Similarity 82.2%; Pred. No. 1.7e-27;  
 Matches 83; Conservative 7; Mismatches 10; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYVKEGERAPLOSELDAKQAKLSKLELSKDIDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDESDSYVKEGLRVPLOSELDAKQAKLSKLELSKDIDELDAETAK-LEKDVDF 60  
 QY 60 KNSDGEQAGYLAAREEDLIAKAELEQTEADLKKAHVEPE 100  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELEKTEADLKKAHVEPE 101  
 RESULT 14  
 AAW14571  
 ID AAW14571 standard; protein; 204 AA.  
 AC AAW14571;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Bf1019.  
 OS WO9709994-A1.  
 XX WO9709994-A1.  
 XX 20-MAR-1997.  
 XX 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-5  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGERAP.....KKAEELEQTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/6CTUS COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	490	100.0	100	4	US-09-147-875A-5
2	472.5	96.4	101	2	US-08-710-749-4
3	465	94.9	100	4	US-09-147-875A-4
4	463	94.5	194	4	US-08-529-055-64
5	463	94.5	8991	4	US-08-714-741-32
6	456	93.1	100	4	US-09-147-875A-6
7	454.5	92.8	101	2	US-08-710-749-3
8	450	91.8	183	4	US-08-529-055-50
9	448	91.4	168	4	US-08-529-055-55
10	447	91.2	98	4	US-09-147-875A-1
11	445.5	90.9	101	2	US-08-710-749-5
12	445	90.8	100	4	US-09-147-875A-2
13	438	89.4	550	4	US-09-583-110-4871
14	438	89.4	550	4	US-09-107-433-3858
15	435	88.8	100	4	US-09-147-875A-3
16	429.5	87.7	99	2	US-08-710-749-9
17	427.5	87.2	101	2	US-08-710-749-1
18	420.5	85.8	101	2	US-08-710-749-2
19	410	83.7	166	4	US-08-529-055-48
20	399.5	81.5	101	4	US-09-147-875A-9
21	397.5	81.1	101	2	US-08-710-749-6
22	395	80.6	100	4	US-09-147-875A-7
23	393	80.2	100	4	US-09-147-875A-8
24	392.5	80.1	185	4	US-08-529-055-46
25	382	78.0	102	2	US-08-710-749-8
26	380.5	77.7	101	2	US-08-710-749-7
27	334.5	68.3	99	2	US-08-710-749-10

28	334.5	68.3	99	4	US-09-147-875A-11	Sequence 11, Appl
29	334.5	68.3	204	4	US-08-529-055-51	Sequence 51, Appl
30	326	66.5	100	4	US-09-147-875A-12	Sequence 12, Appl
31	322.5	65.8	99	4	US-09-147-875A-16	Sequence 16, Appl
32	318.5	65.0	99	2	US-08-710-749-11	Sequence 11, Appl
33	318.5	65.0	198	4	US-08-529-055-61	Sequence 61, Appl
34	318.5	65.0	619	1	US-08-465-746-2	Sequence 2, Appl
35	318.5	65.0	619	1	US-08-214-164-2	Sequence 2, Appl
36	318.5	65.0	619	2	US-08-467-852A-3	Sequence 3, Appl
37	318.5	65.0	619	2	US-08-246-636-2	Sequence 2, Appl
38	318.5	65.0	619	2	US-08-247-491A-3	Sequence 3, Appl
39	318.5	65.0	619	2	US-08-319-795-2	Sequence 2, Appl
40	318.5	65.0	619	2	US-08-468-985-2	Sequence 2, Appl
41	318.5	65.0	619	3	US-08-312-949-2	Sequence 2, Appl
42	318.5	65.0	648	1	US-08-072-070-2	Sequence 2, Appl
43	318.5	65.0	648	1	US-08-469-434-2	Sequence 2, Appl
44	318.5	65.0	648	1	US-08-214-222-2	Sequence 2, Appl
45	318.5	65.0	648	2	US-08-467-852A-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-147-875A-5  
; Sequence 5, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-5

Query Match 100.0%; Score 490; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 8.5e-39;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDSESDYVKEGERAPLQSELDQAQKLSLEESDKIDELDAETAKLEKQVEDPK 60  
DB 1 LKEIDSESDYVKEGERAPLQSELDQAQKLSLEESDKIDELDAETAKLEKQVEDPK 60  
QY 61 NSDGEQAGQYLAARAEEDLIAKKAELEQTEADLKAVHEPE 100  
DB 61 NSDGEQAGQYLAARAEEDLIAKKAELEQTEADLKAVHEPE 100

RESULT 2  
US-08-710-749-4  
; Sequence 4, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/710,749  
;; FILING DATE: 20-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 101 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
US-08-710-749-4

Query Match 96.4%; Score 472.5; DB 2; Length 101;  
Best Local Similarity 98.0%; Pred. No. 3.7e-37;  
Matches 99; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
  
Qy 1 LKEIDESSEYVKEGERAPLQSEL-DAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 59  
|||||  
Db 1 LKEIDESSEYVKEGERAPLQSELDDAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 60  
|||||  
  
Qy 60 KNSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEP 100  
|||||  
Db 61 KNSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVNEP 101  
|||||

RESULT 3  
US-09-147-875A-4  
; Sequence 4, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-4

Query Match 94.9%; Score 465; DB 4; Length 100;  
Best Local Similarity 96.0%; Pred. No. 1.8e-36;  
Matches 96; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 LKEIDESSEYVKEGERAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 60  
|||||  
Db 1 LKEIDESSEYVKEGERAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 60  
|||||  
  
Qy 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEP 100  
|||||  
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVDEP 100  
|||||

RESULT 4  
US-08-529-055-64  
; Sequence 64, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:

;; APPLICANT: Briles, David B.  
;; APPLICANT: McDaniel, Larry S.  
;; APPLICANT: Swiatlo, Edwin  
;; APPLICANT: Yother, Janet  
;; APPLICANT: Brooks-Walter, Alexis  
;; TITLE OF INVENTION: Pneumococcal Genes, Portions  
;; TITLE OF INVENTION: Thereof, Expression Products  
;; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
;; TITLE OF INVENTION: Portions and Products  
;; NUMBER OF SEQUENCES: 73  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Curtis, Morris & Safford, P.C.  
;; STREET: 530 Fifth Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/529,055  
;; FILING DATE: 15-SEP-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2400  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 64:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 194 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-529-055-64

Query Match 94.5%; Score 463; DB 4; Length 194;  
Best Local Similarity 97.0%; Pred. No. 6.2e-36;  
Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 LKEIDESSEYVKEGERAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 60  
|||||  
Db 1 LKEIDESSEYVKEGERAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKQVEDF 60  
|||||

Qy 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEP 99  
|||||  
Db 61 XSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVNEP 99  
|||||

RESULT 5  
US-08-714-741-32  
; Sequence 32, Application US/08714741  
; Patent No. 6500613  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David B.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
; TITLE OF INVENTION: PORTIONS AND PRODUCTS  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford, P.C.

STREET: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: U.S.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,741

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer Esq., William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2460

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 8991 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-08-714-741-32

Query Match

Best Local Similarity 94.5%; Score 463; DB 4; Length 8991;

Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60

DB 7537 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 7596

QY 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 99

DB 7597 XSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHP 7635

RESULT 6

US-09-147-875A-6

Sequence 6, Application US/09147875A

Patent No. 6638516

GENERAL INFORMATION:

APPLICANT: BECKER et al.

TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

FILE REFERENCE: 454312-2471

CURRENT APPLICATION NUMBER: US/09/147,875A

CURRENT FILING DATE: 1999-05-24

NUMBER OF SEQ ID NOS: 28

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 6

LENGTH: 100

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: UNSURE

LOCATION: (1)..(100)

OTHER INFORMATION: amino acid 'Xaa' can be any amino acid

US-09-147-875A-6

Query Match

Best Local Similarity 93.1%; Score 456; DB 4; Length 100;

Matches 94; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60

DB 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLELSKIDELDAETAKLEKVEDPK 60

QY 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHEPE 100

DB 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHEPE 100

RESULT 7

US-08-710-749-3

Sequence 3, Application US/08710749

Patent No. 5955089

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: Hollingshead, Susan

APPLICANT: Becker, Robert

TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE

TITLE OF INVENTION: PROTEINS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

STREET: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/710,749

FILING DATE: 20-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2074

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 101 amino acids

TYPE: amino acid

STRANDEDNESS: n/a

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-08-710-749-3

Query Match

Best Local Similarity 92.8%; Score 454.5; DB 2; Length 101;

Matches 96; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGERAPLOSEL-DAKQAKLSKLELSKIDELDAETAKLEKVEDP 59

DB 1 LKEIDSDSDYVKEGERAPLOSELDDAKQAKLSKLELSKIDELDAETAKLEKVEDP 60

QY 60 KNSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHEPE 100

DB 61 KNSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVHEPE 101

RESULT 8

US-08-529-055-50

Sequence 50, Application US/08529055

Patent No. 6592876

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: McDaniel, Larry S.

APPLICANT: Swiatlo, Edwin

APPLICANT: Yother, Janet

APPLICANT: Brooks-Walter, Alexis

TITLE OF INVENTION: Pneumococcal Genes, Portions





APPLICANT: Hollingshead, Susan  
APPLICANT: Becker, Robert  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-5

Query Match 90.9%; Score 445.5; DB 2; Length 101;  
Best Local Similarity 93.1%; Pred. No. 1.2e-34;  
Matches 94; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGERAPLQSEL-DAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 59  
DB 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 60  
QY 60 KNSDGEQAGYLAAREDLIAKAELEQTEADLKKAHVEPE 100  
DB 61 KNSDGEQAGYLAAREDLIAKAELEQTEADLKKAHVEPE 101

RESULT 12  
US-09-147-875A-2  
Sequence 2, Application US/09147875A  
Patent No. 6638516  
GENERAL INFORMATION:  
APPLICANT: BECKER et al.  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
FILE REFERENCE: 454312-2471  
CURRENT APPLICATION NUMBER: US/09/147,875A  
CURRENT FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-2

Query Match 90.8%; Score 445; DB 4; Length 100;  
Best Local Similarity 92.0%; Pred. No. 1.4e-34;  
Matches 92; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 60

DB 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 60  
QY 61 NSDGEQAGYLAAREDLIAKAELEQTEADLKKAHVEPE 100  
DB 61 NSNGEQAEYRAAREDLIAKAELEQTEADLKKAHVEPE 100  
RESULT 13  
US-09-583-110-4871  
Sequence 4871, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
FILE REFERENCE: PATH00-07A  
CURRENT APPLICATION NUMBER: US/09/583,110  
CURRENT FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 4871  
LENGTH: 550  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4871

Query Match 89.4%; Score 438; DB 4; Length 550;  
Best Local Similarity 91.0%; Pred. No. 4.6e-33;  
Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 60  
DB 144 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVDF 203  
QY 61 NSDGEQAGYLAAREDLIAKAELEQTEADLKKAHVEPE 100  
DB 204 NSNGEQAEYRAAREDLIAKAELEQTEADLKKAHVEPE 243

RESULT 14  
US-09-107-433-3858  
Sequence 3858, Application US/09107433  
Patent No. 6800744  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bugh  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
THERAPEUTICS  
NUMBER OF SEQUENCES: 5206  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: <Unknown>  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: <Unknown>  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,433  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/ 085131  
FILING DATE: May 12, 1998

```

; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match      89.4%; Score 438; DB 4; Length 550;
Best Local Similarity 91.0%; Pred. No. 4.6e-33;
Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 144 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 203
   |||||

QY 61 NSDGEQAGQYLAAREEDLIAKAELEQTEADLKKAVHEPE 100
   |||||
Db 204 NSNGEQARQYRAAREEDLAAKQAELEKTEADLKKAVHEPE 243
   |||||

RESULT 15
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

Query Match      88.8%; Score 435; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 1.2e-33;
Matches 90; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||
Db 1 LKEIDSESDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
   |||||

QY 61 NSDGEQAGQYLAAREEDLIAKAELEQTEADLKKAVHEPE 100
   |||||
Db 61 NSNGEQARQYRAAREEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||

```

Search completed: June 18, 2005, 17:07:06  
Job time : 19.9189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-5  
Perfect score: 490  
Sequence: 1 LKEIDSDSDYKGERAP.....KKALEQTEADLKKAHVHEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	490	100.0	100	15	US-10-674-755-5
2	465	94.9	100	15	US-10-674-755-4
3	463	94.5	194	15	US-10-299-636-79
4	456	93.1	100	15	US-10-674-755-6
5	450	91.8	183	15	US-10-299-636-65
6	448	91.4	168	15	US-10-299-636-70
7	447	91.2	98	15	US-10-674-755-1
8	445	90.8	100	15	US-10-674-755-2
9	435	88.8	100	15	US-10-674-755-3
10	410	83.7	166	15	US-10-299-636-63
11	399.5	81.5	101	15	US-10-674-755-9
					Sequence 5, Appli
					Sequence 4, Appli
					Sequence 79, Appli
					Sequence 6, Appli
					Sequence 65, Appli
					Sequence 70, Appli
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 63, Appli
					Sequence 9, Appli

12	395	80.6	100	15	US-10-674-755-7	Sequence 7, Appli
13	393	80.2	100	15	US-10-674-755-8	Sequence 8, Appli
14	392.5	80.1	185	15	US-10-299-636-61	Sequence 61, Appli
15	394.5	68.3	99	15	US-10-674-755-11	Sequence 11, Appli
16	334.5	68.3	204	15	US-10-299-636-66	Sequence 66, Appli
17	326	66.5	100	15	US-10-674-755-12	Sequence 12, Appli
18	322.5	65.8	99	15	US-10-674-755-16	Sequence 16, Appli
19	318.5	65.0	198	15	US-10-299-636-76	Sequence 76, Appli
20	318.5	65.0	354	15	US-10-299-636-105	Sequence 105, Appli
21	318.5	65.0	588	15	US-10-299-636-96	Sequence 96, Appli
22	318.5	65.0	619	10	US-09-882-774-1	Sequence 1, Appli
23	318.5	65.0	619	15	US-10-282-122A-73702	Sequence 73702, A
24	318.5	65.0	619	16	US-10-414-532-72	Sequence 72, Appli
25	314.5	64.2	170	15	US-10-299-636-75	Sequence 75, Appli
26	314.5	64.2	181	15	US-10-299-636-57	Sequence 57, Appli
27	314.5	64.2	643	15	US-10-299-636-95	Sequence 95, Appli
28	314.5	64.2	670	9	US-09-748-875-63	Sequence 63, Appli
29	314.5	64.2	670	10	US-09-298-523B-63	Sequence 63, Appli
30	314.5	64.2	690	9	US-09-748-875-61	Sequence 61, Appli
31	314.5	64.2	690	10	US-09-298-523B-61	Sequence 61, Appli
32	314.5	64.2	691	9	US-09-748-875-1	Sequence 1, Appli
33	314.5	64.2	691	10	US-09-298-523B-1	Sequence 1, Appli
34	314.5	64.2	701	9	US-09-748-875-62	Sequence 62, Appli
35	314.5	64.2	701	10	US-09-298-523B-62	Sequence 62, Appli
36	314.5	64.2	707	9	US-09-748-875-2	Sequence 2, Appli
37	314.5	64.2	707	10	US-09-298-523B-2	Sequence 2, Appli
38	314.5	64.2	711	9	US-09-748-875-3	Sequence 3, Appli
39	314.5	64.2	711	10	US-09-298-523B-3	Sequence 3, Appli
40	314.5	64.2	739	17	US-10-732-923-3294	Sequence 3294, Ap
41	314.5	64.2	929	9	US-09-748-875-60	Sequence 60, Appli
42	314.5	64.2	929	10	US-09-298-523B-60	Sequence 60, Appli
43	314.5	64.2	929	15	US-10-299-636-94	Sequence 94, Appli
44	313.5	64.0	99	15	US-10-674-755-13	Sequence 13, Appli
45	311.5	63.6	188	15	US-10-299-636-74	Sequence 74, Appli

ALIGNMENTS

RESULT 1  
US-10-674-755-5  
; Sequence 5, Application US/10674755  
; Publication NO. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-5

Query Match 100.0%; Score 490; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.3e-32;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYKGERAPLOSELDAKQAKLSLEELSDKIDELDAETAKLEKVEDPK 60  
DB 1 LKEIDSDSDYKGERAPLOSELDAKQAKLSLEELSDKIDELDAETAKLEKVEDPK 60  
QY 61 NSDGEQAGQYLAAREEDLIKKALEQTEADLKKAHVHEPE 100  
DB 61 NSDGEQAGQYLAAREEDLIKKALEQTEADLKKAHVHEPE 100  
RESULT 2

```
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match          94.9%; Score 465; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 4.3e-30;
Matches 96; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60
Db 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60

Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 100
Db 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 100

RESULT 3
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match          94.5%; Score 463; DB 15; Length 194;
Best Local Similarity 97.0%; Pred. No. 1.3e-29;
Matches 96; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60
Db 1 LKEIDSESDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60
```

```
Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 99
Db 61 XSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 99

RESULT 4
US-10-674-755-6
; Sequence 6, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (1)...(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match          93.1%; Score 456; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 2.3e-29;
Matches 94; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60
Db 1 LKEIDSESDYVKEGERAPLQSELDKQAKLSKLEELSDKIDELDAEIAKLEKQVDFK 60

Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 100
Db 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVPE 100

RESULT 5
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match          91.8%; Score 450; DB 15; Length 183;
Best Local Similarity 93.0%; Pred. No. 1.3e-28;
```

Matches 93; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
Db 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100  
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100

RESULT 6

US-10-299-636-70  
; Sequence 70, Application US/10299636  
; Publication No. US2004007847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa at position 38 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (42)  
; OTHER INFORMATION: Xaa at position 42 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (84)  
; OTHER INFORMATION: Xaa at position 84 is unknown  
US-10-299-636-70

Query Match 91.4%; Score 448; DB 15; Length 168;

Best Local Similarity 93.0%; Pred. No. 1.8e-28;  
Matches 93; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
Db 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100  
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100

RESULT 7

US-10-674-755-1  
; Sequence 1, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-1

Query Match 91.2%; Score 447; DB 15; Length 98;  
Best Local Similarity 96.0%; Pred. No. 1.2e-28;  
Matches 96; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
Db 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100  
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 98

RESULT 8

US-10-674-755-2  
; Sequence 2, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-2

Query Match 90.8%; Score 445; DB 15; Length 100;

Best Local Similarity 92.0%; Pred. No. 1.7e-28;  
Matches 92; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
Db 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKQVEDFK 60  
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100  
Db 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKKAVHEPE 100

RESULT 9

US-10-674-755-3  
; Sequence 3, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3

```
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      88.8%; Score 435; DB 15; Length 100;
Best Local Similarity 90.0%; Pred. No. 1.1e-27;
Matches 90; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSDGEAQGYLAAAEEDLIAKAELEQTEADLKKAVHEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSNGEAEQYRAAGEDLAAKQAELEKTEADLKKAVHEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      83.7%; Score 410; DB 15; Length 166;
Best Local Similarity 85.0%; Pred. No. 1.9e-25;
Matches 85; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSDGEAQGYLAAAEEDLIAKAELEQTEADLKKAVHEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEAQGYLVAEAKDLDAKEAEELGNTGADLKKAVDEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      81.5%; Score 399.5; DB 15; Length 101;
Best Local Similarity 83.2%; Pred. No. 7.9e-25;
Matches 84; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 60 KNSDGEAQGYLAAAEEDLIAKAELEQTEADLKKAVHEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-10-674-755-7
; Sequence 7, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-7

Query Match      80.6%; Score 395; DB 15; Length 100;
Best Local Similarity 83.0%; Pred. No. 1.8e-24;
Matches 83; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKEIDESDYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKDVDFK 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 NSDGEAQGYLAAAEEDLIAKAELEQTEADLKKAVHEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSDGEAQGYLVAEAKDLDAKEAEELGNTGADLKKAVDEPE 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-5  
Perfect score: 490  
Sequence: 1 LKEIDSESDYVKEGERAP.....KKAELEQTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	318.5	65.0	619	2 A97887	surface protein ps
2	318.5	65.0	619	2 A41971	surface protein ps
3	126	25.7	744	2 F95013	pneumococcal surfa
4	116	23.7	886	2 H63378	conserved hypotet
5	111.5	22.8	764	2 T05409	hypothetical prote
6	111.5	22.8	896	2 S43074	epidermal growth f
7	110	22.4	1269	2 F84730	probable myosin he
8	108	22.0	1169	2 A64505	P115 homolog - Met
9	103.5	21.1	897	2 A54696	EGF receptor subsc
10	103	21.0	1190	2 E84193	chromosome segrega
11	102	20.8	501	2 A44643	M protein precursor
12	102	20.8	1156	2 B70356	chromosome assembl
13	101.5	20.7	522	2 G02533	occludin - human
14	101	20.6	405	2 A33939	Fc gamma (IgG) rec
15	101	20.6	1319	2 A28313	glued protein - fr
16	100.5	20.5	1027	2 S37711	kinesin heavy chai
17	100	20.4	388	2 A46173	Mrp4 protein - Str
18	99.5	20.3	1053	2 A41642	dynactin - chicken
19	99.5	20.3	1110	2 I51116	NF-180 - sea lamp
20	99	20.2	387	2 S57834	fcrA protein precu
21	99	20.2	852	2 D72230	conserved hypotet
22	99	20.2	1006	2 C70445	ATPase subunit of
23	99	20.2	1179	2 G95144	conserved hypotet
24	98	20.0	3488	2 T34418	hypothetical prote
25	97.5	19.9	1116	2 D97001	probable membrane
26	97	19.8	646	2 AD3409	ABC transporter AT
27	97	19.8	911	2 S51441	hypothetical prote
28	96.5	19.7	924	2 S06117	myosin heavy chain
29	96.5	19.7	1177	2 B75150	chromosome segrega

ALIGNMENTS

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID:  
C:Genetics:  
A:Gene: pspA

Query Match

Best Local Similarity 65.0%; Score 318.5; DB 2; Length 619;

Matches 71; Conservative 10; Mismatches 9; Indels 19; Gaps 2;

QY 1 LKEIDSESDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLE----- 53

DB 223 LKEIDSESDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLE----- 282

QY 54 --KDVEDFKNSDGEQAGOVLAABEDLIATKAAELQTEADLKAVHEPE 100

DB 283 ENNVEDY-----FKGLEKTIATKAAELQTEADLKAVHEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A>Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:G153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991

A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability

A;Reference number: A60282; MUID:91169598; PMID:2004810

A;Accession: A60282

A:Molecule type: protein

A;Residues: 32-76 <TAI>

A;Experimental source: strain JY2008

C;Genetics:

A;Gene: pspA

F;1-31/Domain: signal sequence #status predicted <SIG>

F;32-619/Product: surface protein pspA #status predicted <MAT>

F;411-430/Domain: cpl repeat homology <CP01>

F;431-450/Domain: cpl repeat homology <CP02>

F;451-470/Domain: cpl repeat homology <CP03>

F;471-490/Domain: cpl repeat homology <CP04>

F;491-510/Domain: cpl repeat homology <CP05>

F;511-530/Domain: cpl repeat homology <CP06>

F;531-550/Domain: cpl repeat homology <CP07>

F;551-570/Domain: cpl repeat homology <CP08>

F;571-591/Domain: cpl repeat homology <CP09>

F;592-611/Domain: cpl repeat homology <CP10>

Query Match 65.0%; Score 318.5; DB 2; Length 619;

Best Local Similarity 65.1%; Pred. No. 4.2e-14;

Matches 71; Conservative 10; Mismatches 9; Indels 19; Gaps 2;

Qy 1 LKEIDESDSYVKEGERAPQSELDKQAKLSKLEELSDKIDELDAEIAKLE----- 53

Db 223 LKEDISESESYAKEGFRAPQSKLDKAKKLSKLEELSDKIDELDAEIAKLEQLKAAE 282

Qy 54 --KQVEPKNSDQAGQAGYLAAREEDLIKAELEQTEADLKKAHVHE 100

Db 283 ENNVEDY-----FKEGLEKTIAAKAELEKTEADLKKAHVHE 321

RESULT 3

F95013

pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C;Accession: F95013

R;Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple, nson, T.; Rickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95013

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-744 <KUR>

A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:C

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP0117

Query Match 25.7%; Score 126; DB 2; Length 744;

Best Local Similarity 32.6%; Pred. No. 0.26;

Matches 43; Conservative 21; Mismatches 28; Indels 40; Gaps 6;

Qy 2 KEIDB-----SDSDYVKEGERAPQSELDKQAKLSK-----LEELSDKI-----D 43

Db 314 KEISNLEILLGGADPED-----DTAALQNLKAAKAELEKQTELEKILDSLDPEGKTQD 368

Qy 44 ELD--AEIAKLEKQVEDPKNS-----DGEQAGYLAAREEDLIKAELE 86

Db 369 ELDKAEAEELDKADELQNLKVADLEKISNLEILLGGADSEDDTAALQNLKATKAELE 428

Qy 87 QTEADLKKAHVHE 98

Db 429 KTKELDAALNE 440

us-10-674-755-5.rpr

RESULT 4

H69378

conserved hypothetical protein AF1032 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69378

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Arttich, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB9021

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 23.7%; Score 116; DB 2; Length 886;

Best Local Similarity 29.1%; Pred. No. 1.4;

Matches 39; Conservative 22; Mismatches 37; Indels 36; Gaps 4;

Qy 1 LKEIDESDSYVKEG-----ERAPQSELDKQAKLSKLEELSDKIDELDAEIAKLEKD- 55

Db 296 LSEINQALRDVKEKREGDLTREAAQIQALKKAEEDNSKLEETIKRIELELELELELEFEKSH 355

Qy 56 --VEDEKNSDGEQAG-----QYLAAABED-----LIAKAE 84

Db 356 RLLETLPKMDRMQGIKAKLEEKNLTPDKVKMYDLLSKAKEEKEITEKLKLIKAKSS 415

Qy 85 LEQTEADLKKAHVHE 98

Db 416 LKTRGAQLKKAHVE 429

RESULT 5

T05409

hypothetical protein F10M6.170 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T05409

R;Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N submitted to the Protein Sequence Database, February 1998

A;Reference number: Z15414

A;Accession: T05409

A:Molecule type: DNA

A;Residues: 1-764 <BEV>

A;Cross-references: UNIPROT:O49371; EMBL:AL021811

A;Experimental source: cultivar Columbia; BAC clone F10M6

C;Genetics:

A;Map position: 4

A;Note: F10M6.170

Query Match 22.8%; Score 111.5; DB 2; Length 764;

Best Local Similarity 32.5%; Pred. No. 2.4;

Matches 37; Conservative 19; Mismatches 41; Indels 17; Gaps 3;

Qy 2 KEIDESDSYVKEGERAPQSELDKQAKLSK-----EELSDKIDELDAEIAKLE 53

Db 163 REIEELKHKLURDERDEAAALQSSLTLLKEELEKMRQEIANKRSKVSWAISFEKSQLLS 222

Qy 54 KDVEDFKNSDGE--QAGYLAAREEDLIKAA-----ELEQTEADLKKAHVHE 98

Db 223 KANENVKQSGEIIYALQRALEKEEELISKATKTKLEKLETEANLKKQTEE 276

RESULT 6

S43074

epidermal growth factor receptor substrate - human



RESULT 12  
B70356  
chromosome assembly protein homolog - Aquifex aeolicus  
C/Species: Aquifex aeolicus  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: B70356  
R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: B70356  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1156 <AQF>  
A/Cross-references: UNIPROT:Q66878; GB:AE000699; NID:G2982328; PIDN:AAC06839.1; PID:G2982  
A/Experimental source: strain VF5  
C/Genetics:  
A/Gene: xcpC  
C/Superfamily: chromosome segregation protein SMC1

Query Match 20.8%; Score 102; DB 2; Length 1156;  
Best Local Similarity 30.0%; Pred. No. 15;  
Matches 30; Conservative 24; Mismatches 38; Indels 8; Gaps 3;  
Qy 1 LKEIDESDSE--DYVKEGERAPLOSELDKAKQAKLSK----EELSDKIDELDAEIAKLEK 54  
Db 854 LQEVKEAEVKKVYDIYIKQEE--LEKEIINLKSGKGLKIKKEELKEKIKFEKKNLKVLEE 911  
Qy 55 DVEDFKNSDGRQAGQGYLAABEDLIAKKAELQTEADLKK 94  
Db 912 KIENLEELKEVEDLKGADDEESIPKLEKLRVTEEIOK 951

RESULT 13  
G02533  
occludin - human  
C/Species: Homo sapiens (man)  
C/Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C/Accession: G02533  
R;Van Itallie, C.M.  
submitted to the EMBL Data Library, April 1996  
A/Reference number: H01403  
A/Accession: G02533  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA  
A/Residues: 1-522 <VAN>  
A/Cross-references: UNIPROT:Q16625; EMBL:U53823; NID:G1322281; PIDN:AAB00195.1; PID:G1322  
C/Superfamily: occludin

Query Match 20.7%; Score 101.5; DB 2; Length 522;  
Best Local Similarity 26.8%; Pred. No. 7.6;  
Matches 30; Conservative 31; Mismatches 24; Indels 27; Gaps 6;  
Qy 5 DESDSEDYVKEGERAPLOSELDKAKQAK----LSKLEELSDKIDELDAEIAKLEKDVED 58  
Db 410 DELS-EDWIR--EYPPITSDQORQLYKKNFDFTGLQYKSLQSELDDEINKELSRDKELDD 466  
Qy 59 FKNSDGEAGQGYLAAAE-----DLIAKKAELQTEADL---KKAV 96  
Db 467 YR----EESSEYMAADEYNRLKQVKSADYKSKKNHCKQLSKLSHKQMV 514

RESULT 14  
A33939  
Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)  
C/Species: Streptococcus sp.  
C/Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 26-Aug-1999  
C/Accession: A33939  
R;Heath, D.G.; Cleary, P.P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989

Search completed: June 18, 2005, 17:03:53  
Job time : 14.113 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-5  
Perfect score: 490  
Sequence: 1 LKIDSESDYVKEGERAP.....KKALEQTEADLKKAHVEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	475	96.9	Q9LAZ0	Q9LAZ0 streptococc
2	464	94.7	Q9LAY6	Q9LAY6 streptococc
3	464	94.7	Q9LAZ1	Q9LAZ1 streptococc
4	460	93.9	Q8KQK5	Q8KQK5 streptococc
5	443	90.4	Q9L591	Q9L591 streptococc
6	438	89.4	Q9L577	Q9L577 streptococc
7	438	89.4	Q9L576	Q9L576 streptococc
8	438	89.4	Q9LAY7	Q9LAY7 streptococc
9	432	88.2	Q9L5B5	Q9L5B5 streptococc
10	432	88.2	Q6UEB2	Q6UEB2 streptococc
11	432	88.2	Q9L568	Q9L568 streptococc
12	432	88.2	Q9L569	Q9L569 streptococc
13	432	88.2	Q9L564	Q9L564 streptococc
14	432	88.2	Q9L567	Q9L567 streptococc
15	432	88.2	Q9L565	Q9L565 streptococc
16	432	88.2	Q9L566	Q9L566 streptococc
17	432	88.2	Q9L570	Q9L570 streptococc
18	432	88.2	Q9L563	Q9L563 streptococc
19	432	88.2	Q9LAZ2	Q9LAZ2 streptococc
20	431	88.0	Q9L581	Q9L581 streptococc
21	431	88.0	Q9L5B6	Q9L5B6 streptococc
22	431	88.0	Q9LAY8	Q9LAY8 streptococc
23	428	87.3	Q9L578	Q9L578 streptococc
24	411	83.9	Q9LAZ3	Q9LAZ3 streptococc
25	400	81.6	Q8GNS9	Q8GNS9 streptococc
26	394	80.4	Q9L592	Q9L592 streptococc
27	394	80.4	Q9LAY9	Q9LAY9 streptococc
28	334.5	68.3	Q9LAY3	Q9LAY3 streptococc
29	318.5	65.0	Q54972	Q54972 streptococc
30	318.5	65.0	Q8DR10	Q8DR10 streptococc
31	314.5	64.2	Q9LAY1	Q9LAY1 streptococc

32	314.5	64.2	739	2	Q9RQT4	Q9RQT4 streptococc
33	314.5	64.2	820	2	Q9RQT1	Q9RQT1 streptococc
34	314.5	64.2	929	2	Q9KK19	Q9KK19 streptococc
35	314.5	64.2	929	2	Q9ZAY5	Q9ZAY5 streptococc
36	302.5	61.7	99	2	Q8KQK4	Q8KQK4 streptococc
37	299.5	61.1	249	2	Q9L575	Q9L575 streptococc
38	298.5	60.9	437	2	Q9LAY4	Q9LAY4 streptococc
39	292.5	59.7	224	2	Q8GNS8	Q8GNS8 streptococc
40	292.5	59.7	426	2	Q9LAY5	Q9LAY5 streptococc
41	287.5	58.7	395	2	Q9LAY2	Q9LAY2 streptococc
42	287.5	58.7	408	2	Q9LAY0	Q9LAY0 streptococc
43	279	56.9	869	2	Q9KK27	Q9KK27 streptococc
44	178.5	36.4	481	2	Q9LAX5	Q9LAX5 streptococc
45	178	36.3	479	2	Q9LAX2	Q9LAX2 streptococc

ALIGNMENTS

RESULT 1  
Q9LAZ0  
ID Q9LAZ0 PRELIMINARY; PRT; 406 AA.  
AC Q9LAZ0;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DL6A;  
RX MEDLINE=20448953; PubMed=10992499;  
DOI=10.1128/JAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071805; AAF27701.1;  
DR InterPro; IPR009082; His\_kin\_Homodim.  
DR InterPro; IPR000533; Tropomyosin.  
FR PRINTS; PR00194; TROPOMYOSIN.  
FT NON TER 406 406  
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 96.9%; Score 475; DB 2; Length 406;  
Best Local Similarity 98.0%; Pred. NO. 7.2e-23;  
Matches 98; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSESDYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKVEDPK 60  
DB 213 LKIDSESDYVKEGERAPLQSELDKAKLSELSKIDELDAETAKLEKVEDPK 272  
QY 61 NSDGEQAGQYLAAREEDLIKKAELQTEADLKKAHVEPE 100  
DB 273 NSDGEQAGQYLAAREEDLIKKAELQTEADLKKAHVEPE 312

RESULT 2  
Q9LAY6  
ID Q9LAY6 PRELIMINARY; PRT; 394 AA.  
AC Q9LAY6;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.

```

OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=L81905;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/JAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RL in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071809; AAF27705.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 394 394
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C88FAA CRC64;

Query Match 94.7%; Score 464; DB 2; Length 394;
Best Local Similarity 96.0%; Pred. No. 3.6e-22;
Matches 96; Conservative 1; Mismatches 23; Indels 0; Gaps 0;

Oy 1 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 60
Db 213 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 272

Oy 61 NSDGEAGQYLAAAEEDLIAKAELEQTEADLKXVHEPE 100
Db 273 NSDGEAGQYLAAAEEDLIAKAELEKAEADLKXVDEPE 312

RESULT 3
O9LAZ1
ID Q9LAZ1 PRELIMINARY; PRT; 395 AA.
AC Q9LAZ1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=BG9739;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/JAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RL in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27700.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECAC41DB7F95 CRC64;

Query Match 94.7%; Score 464; DB 2; Length 395;
Best Local Similarity 96.0%; Pred. No. 3.6e-22;
Matches 96; Conservative 1; Mismatches 23; Indels 0; Gaps 0;

Oy 1 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 60
Db 213 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 272

Oy 61 NSDGEAGQYLAAAEEDLIAKAELEQTEADLKXVHEPE 100
Db 273 NSDGEAGQYLAAAEEDLIAKAELEKAEADLKXVDEPE 312
```

```

RESULT 4
O8KQK5
ID O8KQK5 PRELIMINARY; PRT; 340 AA.
AC O8KQK5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/JAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1 1
FT NON TER 340 340
SQ SEQUENCE 340 AA; 38023 MW; EE07ECF00B1FBD57 CRC64;

Query Match 93.9%; Score 460; DB 2; Length 340;
Best Local Similarity 94.0%; Pred. No. 5.6e-22;
Matches 94; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 60
Db 197 LKEIDESSEYVKEGERAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEKXVEDFK 256

Oy 61 NSDGEAGQYLAAAEEDLIAKAELEQTEADLKXVHEPE 100
Db 257 NSDGEAGQYLAAAEEDLVAKAELEKTEADLKXVNEPE 296

RESULT 5
O9L591
ID O9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B.W.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
```





DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON_TER	1
FT	NON_TER	218
SQ	SEQUENCE	218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;
Query Match	88.2%; Score 432; DB 2; Length 218;	
Best Local Similarity	89.0%; Pred. No. 2.3e-20;	
Matches	89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;	
QY	1 LKEIDSESDYVKEGERAPLQSLDAKQAKLSKLELSDKIDELDAEIAKLEKQVEDFK 60	
DB	27 LKEIDSESDYIIEGLRAPLQSLDKAKKAKLSKLELSDKIDELDAEIAKLEKQVEDFK 86	
QY	61 NSDGEQAGYLAAREEDLIAKKAELQTEADLKKAVHEP 100	
DB	87 NSDGEQAEQYLVAKKDLDAKKAELTEADLKKAVDEP 126	
RESULT 11		
Q9L568	PRELIMINARY; PRT; 233 AA.	
ID	Q9L568	
AC	Q9L568;	
DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)	
DE	FspA (Fragment).	
GN	Name=pspA;	
OS	Streptococcus pneumoniae.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
ON	NCBI_TaxID=1313;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RX	MEDLINE=20472698; PubMed=11015380;	
RA	Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;	
RT	"Pneumococcal pspA sequence types of prevalent multiresistant	
RT	pneumococcal strains in the United States and of internationally	
RT	disseminated clones.";	
RL	J. Clin. Microbiol. 38:3663-3669(2000).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=39;	
RA	Beall B.W.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF265902; AAF70092.1; -	
DR	InterPro; IPR009082; His_kin_homodim.	
FT	NON_TER	1
FT	NON_TER	233
SQ	SEQUENCE	233 AA; 24514 MW; D5C494019C45BFE2 CRC64;
Query Match	88.2%; Score 432; DB 2; Length 233;	
Best Local Similarity	89.0%; Pred. No. 2.4e-20;	
Matches	89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;	
QY	1 LKEIDSESDYVKEGERAPLQSLDAKQAKLSKLELSDKIDELDAEIAKLEKQVEDFK 60	
DB	28 LKEIDSESDYIIEGLRAPLQSLDKAKKAKLSKLELSDKIDELDAEIAKLEKQVEDFK 87	
QY	61 NSDGEQAGYLAAREEDLIAKKAELQTEADLKKAVHEP 100	
DB	88 NSDGEQAEQYLVAKKDLDAKKAELTEADLKKAVDEP 127	
RESULT 12		
Q9L569	PRELIMINARY; PRT; 236 AA.	
ID	Q9L569	
AC	Q9L569;	
DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)	
DE	FspA (Fragment).	
GN	Name=pspA;	

```
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=177;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=177;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255901; AAF70091.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR NON_TER 1
FT NON_TER 236
SQ SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;

Query Match 88.2%; Score 432; DB 2; Length 236;
Best Local Similarity 89.0%; Pred. No. 2.5e-20;
Matches 89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
Db 49 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 108
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEPE 100
Db 109 NSDGEQAEQYLVAKKDLDAKKAELTEADLKAVDEPE 148

RESULT 13
Q9L564
ID Q9L564 PRELIMINARY; PRT; 243 AA.
AC Q9L564;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70096.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 88.2%; Score 432; DB 2; Length 243;
```

```
Best Local Similarity 89.0%; Pred. No. 2.5e-20;
Matches 89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
Db 74 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 133
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEPE 100
Db 134 NSDGEQAEQYLVAKKDLDAKKAELTEADLKAVDEPE 173

RESULT 14
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255903; AAF70093.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match 88.2%; Score 432; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 2.5e-20;
Matches 89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 60
Db 50 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEKQVEDPK 109
QY 61 NSDGEQAGQYLAABEDLIAKAELEQTEADLKAVHEPE 100
Db 110 NSDGEQAEQYLVAKKDLDAKKAELTEADLKAVDEPE 149

RESULT 15
Q9L565
ID Q9L565 PRELIMINARY; PRT; 244 AA.
AC Q9L565;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
```

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=183;  
 RX MEDLINE=20472698; PubMed=11015380;  
 RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;  
 RT "Pneumococcal pspA sequence types of prevalent multiresistant  
 RT pneumococcal strains in the United States and of internationally  
 RT disseminated clones."; J. Clin. Microbiol. 38:3663-3669(2000).  
 RL J. Clin. Microbiol. 38:3663-3669(2000).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=183;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF255905; AAF70095.1; -;  
 DR InterPro; IPR009082; His\_kin\_homodim.  
 FT NON\_TER 1  
 FT NON\_TER 244 244  
 SQ SEQUENCE 244 AA; 25946 MW; F9274FFD1957DD06 CRC64;

Query Match 88.2%; Score 432; DB 2; Length 244;  
 Best Local Similarity 89.0%; Pred. No. 2.5e-20;  
 Matches 89; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLOSELDAKQAKLSKLEELSDKIDELDAEIAKLEKVDK 60  
 Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 51 LKEIDESDSEYIKELRAPLOSKLDKAKKLSKLEELSDKIDELDAEIAKLEKVDK 110  
 Qy 61 NSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAVHEPE 100  
 Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 111 NSDGEQAEQYLVNAKKDLDAKKAELNTEADLKKAVDEPE 150

Search completed: June 18, 2005, 17:01:35  
 Job time : 61.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-6  
Perfect score: 477  
Sequence: 1 LKEIDSESDYKGERAP.....KKXLEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	97.1	168	7	ABW02609
2	463	97.1	8991	6	ABU08487
3	449	94.1	183	2	AAW14570
4	449	94.1	183	7	ABW02604
5	447.5	93.8	167	2	AAW14575
6	438	91.8	194	2	AAW14584
7	438	91.8	194	7	ABW02618
8	421	88.3	550	8	ADK48356
9	421	88.3	550	8	ADR95223
10	396	83.0	166	2	AAW14568
11	396	83.0	166	7	ABW02602
12	375.5	78.7	185	2	AAW14586
13	375.5	78.7	185	7	ABW02600
14	317.5	66.6	204	2	AAW14571
15	317.5	66.6	204	7	ABW02605
16	313.5	65.7	170	7	ABW02614
17	313.5	65.7	181	7	ABW02556
18	313.5	65.7	865	6	ABU08489
19	313.5	65.7	929	2	AAW14593
20	313.5	65.7	929	2	AAW14384
21	310.5	65.1	188	2	AAW14580
22	310.5	65.1	188	7	ABW02613
23	304.5	63.8	198	2	AAW14581
24	301.5	63.2	198	7	ABW02615
25	301.5	63.2	315	2	AAW04375

26	301.5	63.2	588	6	ABU08491	Abu08491 Coiled co
27	301.5	63.2	589	2	AAW43392	Aay43392 PspC alph
28	301.5	63.2	619	2	AAR63437	Aar63437 Pneumococ
29	301.5	63.2	619	2	AAR87598	Aar87598 Pneumococ
30	301.5	63.2	619	2	AAR86911	Aar86911 Pneumococ
31	301.5	63.2	619	2	AAW18338	Aay18338 Streptoco
32	301.5	63.2	619	5	AAE18782	Aae18782 S. pneumo
33	301.5	63.2	619	6	ABU45778	Abu45778 Protein e
34	301.5	63.2	619	8	ADO52126	Ado52126 Streptoco
35	301.5	63.2	648	2	AAW70336	Aaw70336 Pneumococ
36	301.5	63.2	648	2	AAW62274	Aaw62274 Streptoco
37	301.5	63.2	648	2	AAW41837	Aay41837 Streptoco
38	301.5	63.2	648	2	AAW87879	Aaw87879 A pneumoc
39	301.5	63.2	653	2	AAW92456	Aaw92456 S. pneumo
40	301.5	63.2	684	2	AAR73912	Aar73912 Streptoco
41	299.5	62.8	204	2	AAW14578	Aaw14578 Streptoco
42	299.5	62.8	204	7	ABW02612	Abw02612 Rct123c p
43	299	62.7	180	2	AAW14562	Aaw14562 Streptoco
44	296	62.1	187	2	AAW14579	Aaw14579 Streptoco
45	292.5	61.3	195	2	AAW14591	Aaw14591 Streptoco

#### ALIGNMENTS

RESULT 1  
ABW02609  
ID ABW02609 standard; protein; 168 AA.  
XX AC ABW02609;  
XX XX  
DT 12-FEB-2004 (first entry)  
XX XX  
DE L81905c pneumococcal surface protein A (PspA) central region.  
XX XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX OS Unidentified.  
XX XX  
XX Key Location/Qualifiers  
FT Misc-difference 1.168  
FT /note= "Xaa = Unknown amino acid"  
XX XX  
PN US6592876-B1.  
XX PD 15-JUL-2003.  
XX PF 15-SEP-1995; 95US-00529055.  
XX PR 20-APR-1993; 93US-00048896.  
XX PR 06-JUN-1995; 95US-00465746.  
XX PA (UABR-) UAB RES FOUND.  
XX PI Briles DE, Medaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

Example 6; SEQ ID NO 55; 121pp; English.

The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic.





```
Db      60 NSDGEQAGQYLAAAEEDLIAKXAKAEADLKKAVDEPE 99

RESULT 6
AAW14584
ID AAW14584 standard; protein; 194 AA.
XX
AC AAW14584;
XX
AC AAW14584;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Db16.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 61
FT /note= "unidentified amino acid"
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
PD
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain WJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX WPI, 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Db16.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 194 AA;
Query Match 91.8%; Score 438; DB 2; Length 194;
Best Local Similarity 91.9%; Pred. No. 2.2e-34;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 1 LKEIDESDSYVKEGEPAPLQSELDAAQAKLSKLEESDKXDELDAEIAKLEKXVEDFK 60
Db 1 LKEIDESDSYVKEGEPAPLQSELDAAQAKLSKLEESDKIDELDAEIAKLEKXVEDFK 60
Qy 61 NSDGEQAGQYLAAAEEDLIAKXAKAEADLKKAVDEP 99
Db 61 XSDGEQAGQYLAAAEEDLIAKAELEQTEADLKKAVNEP 99
Qy 61 NSDGEQAGQYLAAAEEDLIAKXAKAEADLKKAVDEP 99
Db 61 XSDGEQAGQYLAAAEEDLIAKAELEQTEADLKKAVNEP 99
RESULT 8
ADK48356
ID ADK48356 standard; protein; 550 AA.

ABW02618
ID ABW02618 standard; protein; 194 AA.
XX
AC ABW02618;
XX
DT 12-FEB-2004 (first entry)
XX
DE Db16ac pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 1.194
FH /note= "Xaa = Unknown amino acid"
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
PD
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
PR
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 64; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, or a
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
XX Sequence 194 AA;
Query Match 91.8%; Score 438; DB 7; Length 194;
Best Local Similarity 91.9%; Pred. No. 2.2e-34;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 1 LKEIDESDSYVKEGEPAPLQSELDAAQAKLSKLEESDKXDELDAEIAKLEKXVEDFK 60
Db 1 LKEIDESDSYVKEGEPAPLQSELDAAQAKLSKLEESDKIDELDAEIAKLEKXVEDFK 60
Qy 61 NSDGEQAGQYLAAAEEDLIAKXAKAEADLKKAVDEP 99
Db 61 XSDGEQAGQYLAAAEEDLIAKAELEQTEADLKKAVNEP 99
RESULT 8
ADK48356
ID ADK48356 standard; protein; 550 AA.
```



```
XX AC ADK48356;
XX XX
XX DT 20-MAY-2004 (first entry)
XX XX
XX DE Streptococcus pneumoniae protein, Seq ID No 4871.
XX KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX OS Streptococcus pneumoniae.
XX PN US6699703-B1.
XX PD 02-MAR-2004.
XX XX
XX PF 26-MAY-2000; 2000US-00583110.
XX XX
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX PR 30-JUN-1998; 98US-00107433.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX XX
XX PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewheat CE;
XX XX
XX DR WPI; 2004-212399/20.
XX DR N-PSDB; ADK45695.
XX XX
XX PT New nucleic acid molecules and polypeptides useful for diagnosing,
XX PT preventing and treating pathological conditions resulting from bacterial
XX PT infection, e.g. Streptococcus pneumoniae infection, and in drug
XX PT screening.
XX PS Disclosure; SEQ ID NO 4871; 301pp; English.
XX XX
XX CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
XX CC and polypeptides. The nucleic acids and proteins are useful for
XX CC diagnosing, preventing and treating pathological conditions resulting
XX CC from bacterial infection, such as S. pneumoniae infection. These may also
XX CC be used for drug screening procedures. The present sequence represents a
XX CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
XX CC data for this patent did not appear in the printed specification but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html.
XX XX
XX SQ Sequence 550 AA;
XX
Query Match 88.3%; Score 421; DB 8; Length 550;
Best Local Similarity 88.0%; Pred. No. 3.3e-32;
Matches 88; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEESDKXDELDAETAKLEKVEDPK 60
DB 144 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDPK 203
QY 61 NSDGEQAGQYLAABEDLIAKXLEKAEADLKKAVIDEPE 100
DB 204 NSNGEQAEQYRAAAEDLIAKQAELEKTEADLKKAVIDEPE 243
RESULT 9
ID ADR95223
XX ADR95223 standard; protein; 550 AA.
XX AC ADR95223;
XX XX
XX DT 16-DEC-2004 (first entry)
XX XX
XX DE Novel S. pneumoniae protein sequence, SEQ ID 3858.
XX KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
XX XX bacterial infection.
```

```
OS Streptococcus pneumoniae.
XX XX
XX PN US6800744-B1.
XX XX
XX PD 05-OCT-2004.
XX XX
XX PF 30-JUN-1998; 98US-00107433.
XX PF 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX XX
XX PI Doucette-Stamm LA, Bush D;
XX XX
XX DR WPI; 2004-697205/68.
XX DR N-PSDB; ADR92620.
XX XX
XX PT New isolated nucleic acid encoding a Streptococcus pneumoniae
XX PT polypeptide, useful for diagnosing, preventing and/or treating
XX PT pathological conditions resulting from the bacterial infection.
XX XX
XX PS Disclosure; SEQ ID NO 3858; 151pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising a sequence
XX CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
XX CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
XX CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
XX CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
XX CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
XX CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
XX CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
XX CC hybridisable under high stringency conditions to the nucleotide sequence.
XX CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
XX CC Also included are a recombinant expression vector comprising the isolated
XX CC nucleic acid cited above operably linked to a transcription regulatory
XX CC element, a cell comprising the recombinant expression vector and a probe
XX CC comprising at least 20 consecutive nucleotides of the nucleotide
XX CC sequences as cited above. The methods and compositions of the present
XX CC invention are useful for the diagnosis, prevention and/or treatment of
XX CC pathological conditions resulting from bacterial infection by
XX CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
XX CC otitis media. The present sequence is one of the 2603 disclosed S.
XX CC pneumoniae protein sequences. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.
XX XX
XX SQ Sequence 550 AA;
XX
Query Match 88.3%; Score 421; DB 8; Length 550;
Best Local Similarity 88.0%; Pred. No. 3.3e-32;
Matches 88; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEESDKXDELDAETAKLEKVEDPK 60
DB 144 LKEIDSDSDYVKEGLRAPLQSELDAKQAKLSKLEELSDKIDELDAETAKLEKVEDPK 203
QY 61 NSDGEQAGQYLAABEDLIAKXLEKAEADLKKAVIDEPE 100
DB 204 NSNGEQAEQYRAAAEDLIAKQAELEKTEADLKKAVIDEPE 243
RESULT 10
AAW14568
ID AAW14568 standard; protein; 166 AA.
XX AC AAW14568;
XX XX
XX DT 17-OCT-2003 (revised)
XX DT 28-OCT-1997 (first entry)
XX XX
XX DE Streptococcus pneumoniae PspA central region.
```

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX Streptococcus pneumoniae; strain Bg8743.  
OS WO9709994-A1.  
PN 20-MAR-1997.  
XX 16-SEP-1996; 96WO-US014819.  
XX 15-SEP-1995; 95US-00529055.  
XX (UABR-) UAB RES FOUND.  
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.  
DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
XX in vaccines for protecting animals against S.pneumoniae infection.  
XX Example 6; Fig 13; 296pp; English.  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 166 AA;  
Query Match 83.0%; Score 396; DB 2; Length 166;  
Best Local Similarity 82.0%; Pred. No. 2.1e-30;  
Matches 82; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKLSEKXSDKXDELDAEIAKLEKXVEDFK 60  
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDKAKKLSKLDSELDKIDELDAEIAKLEKXVEDFK 60  
QY 61 NSDGEQAGQYLAAAEEDLIAKXLEKAEADLKKAVIDEPE 100  
DB 61 NSDGEQAGQYLVAEKOLDKAEAEELGNTGADLKKAVIDEPE 100  
RESULT 11  
ABW02602  
ID ABW02602 standard; protein; 166 AA.  
XX  
AC ABW02602;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Bg8743c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX immunological; gene therapy; immunostimulant.  
XX Unidentified.  
OS  
XX US6592876-B1.  
PN  
XX 15-JUL-2003.  
PD

XX 15-SEP-1995; 95US-00529055.  
XX 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX (UABR-) UAB RES FOUND.  
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
PI WPI; 2003-862841/80.  
DR  
XX Immunological composition for obtaining expression products used for  
XX detecting the presence of Streptococcus pneumoniae or its strain,  
XX comprises at least two different full length isolated gene encoding  
XX pneumococcal surface protein A.  
XX Example 6; SEQ ID NO 48; 121pp; English.  
XX The present invention relates to an immunological composition comprising  
XX at least 2 different full length isolated genes encoding pneumococcal  
XX surface protein A (PspAs) from different groups based on restriction  
XX fragment polymorphism analysis. The invention is useful for obtaining  
XX expression products by recombinant techniques to detect, determine,  
XX isolate or diagnose the presence of Streptococcus pneumoniae or its  
XX strain. The expression product is useful for preparing antigenic,  
XX immunological or vaccine compositions, for eliciting antibodies, an  
XX immunological response (other than or additional to antibodies) or a  
XX protective response (including antibody or other immunological response  
XX by administering compositions to a host). The invention is also useful as  
XX vaccines and in gene therapy. The present sequence is Bg8743c  
XX pneumococcal surface protein A (PspA) central region. This sequence is  
XX used in the exemplification of the invention  
XX  
SQ Sequence 166 AA;  
Query Match 83.0%; Score 396; DB 7; Length 166;  
Best Local Similarity 82.0%; Pred. No. 2.1e-30;  
Matches 82; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 1 LKEIDESDSEYVKEGERAPLQSELDKAKLSEKXSDKXDELDAEIAKLEKXVEDFK 60  
DB 1 LKEIDESDSEYVKEGLRAPLQSKLDKAKKLSKLDSELDKIDELDAEIAKLEKXVEDFK 60  
QY 61 NSDGEQAGQYLAAAEEDLIAKXLEKAEADLKKAVIDEPE 100  
DB 61 NSDGEQAGQYLVAEKOLDKAEAEELGNTGADLKKAVIDEPE 100  
RESULT 12  
AAW14566  
ID AAW14566 standard; protein; 185 AA.  
XX  
AC AAW14566;  
XX  
DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
DE Streptococcus pneumoniae PspA central region.  
XX  
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX Streptococcus pneumoniae; strain Ac94.  
OS  
XX WO9709994-A1.  
PN  
XX 20-MAR-1997.  
XX  
XX 16-SEP-1996; 96WO-US014819.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 78.7%; Score 375.5; DB 2; Length 185;  
 Best Local Similarity 79.2%; Pred. No. 2.3e-28;  
 Matches 80; Conservative 6; Mismatches 14; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSEDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDSDSEDYVKEGLRVPQLQSELDVQAKLLKLEELSDKIDELDAETAKLKKDVDF 60  
 QY 60 KNSDGEQAGQYLAAREEDLIKAKLEKAEADLKKAVIDEPE 100  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELEKTEADLKKAVIDEPE 101  
 RESULT 13  
 AEW02600 standard; protein; 185 AA.  
 AC AEW02600;  
 XX 12-FEB-2004 (first entry)  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 OS US6592876-B1.  
 PN US6592876-B1.  
 XX 15-JUL-2003.  
 PD 15-SEP-1995; 95US-00529055.  
 PF 15-SEP-1995; 95US-00529055.  
 XX 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX (UABR-) UAB RES FOUND.  
 PA Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR Immunological composition for obtaining expression products used for  
 XX detecting the presence of Streptococcus pneumoniae or its strain,  
 PT

PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 46; 121pp; English.  
 PS The present invention relates to an immunological composition comprising  
 XX at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 78.7%; Score 375.5; DB 7; Length 185;  
 Best Local Similarity 79.2%; Pred. No. 2.3e-28;  
 Matches 80; Conservative 6; Mismatches 14; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSEDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAETAK-LEKDVDF 59  
 DB 1 LKEIDSDSEDYVKEGLRVPQLQSELDVQAKLLKLEELSDKIDELDAETAKLKKDVDF 60  
 QY 60 KNSDGEQAGQYLAAREEDLIKAKLEKAEADLKKAVIDEPE 100  
 DB 61 QNSGGGYSALYLEAEKDLVAKKAELEKTEADLKKAVIDEPE 101  
 RESULT 14  
 AAW14571  
 ID AAW14571 standard; protein; 204 AA.  
 XX AAW14571;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Ef1019.  
 OS WO9709994-A1.  
 PN 20-MAR-1997.  
 PD 16-SEP-1996; 96WO-US014819.  
 PF 15-SEP-1995; 95US-00529055.  
 PR (UABR-) UAB RES FOUND.  
 PA Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 204 AA;

Query Match 66.6%; Score 317.5; DB 2; Length 204;  
Best Local Similarity 66.1%; Pred. No. 1e-22;  
Matches 72; Conservative 6; Mismatches 12; Indels 19; Gaps 2;  
QY 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEEXSKXDELDAEIAKLE----- 53  
DB 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLELSKIDELDAEIAKLEPDLKAAE 60  
QY 54 --KDVEDFKNSDGEQAGQYLAARAEEDLIAXKXLEKAEADLKXAVDEPE 100  
DB 61 ENNVEDY-----FKEGLEKTIAAKKAELKTEADLKXAVNEPE 99

#### RESULT 15

ABW02605  
ID ABW02605 standard; protein; 204 AA.

AC ABW02605;

XX 12-FEB-2004 (first entry)

XX Ef1019c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 51; 121pp; English.

XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Ef1019c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention

XX Sequence 204 AA;

Query Match 66.6%; Score 317.5; DB 7; Length 204;  
Best Local Similarity 66.1%; Pred. No. 1e-22;  
Matches 72; Conservative 6; Mismatches 12; Indels 19; Gaps 2;

QY 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLEEXSKXDELDAEIAKLE----- 53  
DB 1 LKEIDSDSDYVKEGERAPLQSELDAKQAKLSKLELSKIDELDAEIAKLEPDLKAAE 60  
QY 54 --KDVEDFKNSDGEQAGQYLAARAEEDLIAXKXLEKAEADLKXAVDEPE 100  
DB 61 ENNVEDY-----FKEGLEKTIAAKKAELKTEADLKXAVNEPE 99

Search completed: June 18, 2005, 16:51:21

Job time : 74.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-6  
Perfect score: 477  
Sequence: 1 LKEIDSESDYVKEGERAP.....KKAXLEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pcp:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	471	98.7	100	4	US-09-147-875A-6
2	464	97.3	100	4	US-09-147-875A-4
3	463	97.1	168	4	US-08-529-055-55
4	463	97.1	8991	4	US-08-714-741-32
5	460.5	96.5	101	2	US-08-710-749-5
6	456	95.6	100	4	US-09-147-875A-5
7	453.5	95.1	101	2	US-08-710-749-3
8	449	94.1	183	4	US-08-529-055-50
9	447.5	93.8	101	2	US-08-710-749-4
10	438	91.8	194	4	US-08-529-055-64
11	421	88.3	98	4	US-09-147-875A-1
12	421	88.3	550	4	US-09-583-110-4871
13	421	88.3	550	4	US-09-107-433-3858
14	419	87.8	100	4	US-09-147-875A-2
15	412.5	86.5	99	2	US-08-710-749-9
16	410.5	86.1	101	2	US-08-710-749-1
17	409	85.7	100	4	US-09-147-875A-3
18	403.5	84.6	101	2	US-08-710-749-2
19	396	83.0	166	4	US-08-529-055-48
20	393.5	82.5	101	2	US-08-710-749-6
21	391	82.0	100	4	US-09-147-875A-7
22	375.5	78.7	185	4	US-08-529-055-46
23	373.5	78.3	101	4	US-09-147-875A-9
24	367	76.9	100	4	US-09-147-875A-8
25	365	76.5	102	2	US-08-710-749-8
26	363.5	76.2	101	2	US-08-710-749-7
27	321.5	67.4	99	4	US-09-147-875A-16

28	317.5	66.6	99	2	US-08-710-749-10	Sequence 10, Appl
29	317.5	66.6	99	4	US-09-147-875A-11	Sequence 11, Appl
30	317.5	66.6	204	4	US-08-529-055-51	Sequence 51, Appl
31	313.5	65.7	170	4	US-08-529-055-60	Sequence 60, Appl
32	313.5	65.7	181	4	US-08-529-055-42	Sequence 42, Appl
33	313.5	65.7	864	4	US-08-714-741-40	Sequence 40, Appl
34	310.5	65.1	99	2	US-08-710-749-17	Sequence 17, Appl
35	310.5	65.1	188	4	US-08-529-055-59	Sequence 59, Appl
36	310	65.0	100	4	US-09-147-875A-10	Sequence 10, Appl
37	309	64.8	100	4	US-09-147-875A-12	Sequence 12, Appl
38	308.5	64.7	99	2	US-08-710-749-15	Sequence 15, Appl
39	301.5	63.2	99	2	US-08-710-749-11	Sequence 11, Appl
40	301.5	63.2	141	4	US-09-286-981B-2	Sequence 2, Appl
41	301.5	63.2	198	4	US-08-529-055-61	Sequence 61, Appl
42	301.5	63.2	588	4	US-08-714-741-42	Sequence 42, Appl
43	301.5	63.2	619	1	US-08-465-746-2	Sequence 2, Appl
44	301.5	63.2	619	1	US-08-214-164-2	Sequence 2, Appl
45	301.5	63.2	619	2	US-08-467-852A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-09-147-875A-6  
; Sequence 6, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-09-147-875A-6

Query Match	98.7%	Score 471;	DB 4;	Length 100;
Best Local Similarity	100.0%;	Pred. No. 4.8e-39;	Mismatches 0;	Indels 0;
Matches	100;	Conservative 0;	0;	Gaps 0;
QY	1	LKEIDSESDYVKEGERAPLOSELDKQAKLSKLEESDKKXDELDAEIAKLEKQVDFK	60	
Db	1	LKEIDSESDYVKEGERAPLOSELDKQAKLSKLEESDKKXDELDAEIAKLEKQVDFK	60	
QY	61	NSDGEQAGQYLAABEDLIAKXLEKAEADLKKAVIDEPE	100	
Db	61	NSDGEQAGQYLAABEDLIAKXLEKAEADLKKAVIDEPE	100	

RESULT 2  
US-09-147-875A-4  
; Sequence 4, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 100  
; TYPE: PRT

```
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

Query Match      97.3%; Score 464; DB 4; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.3e-38;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||
Db 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||
Db 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||

RESULT 3
US-08-529-055-55
; Sequence 55, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-55

Query Match      97.1%; Score 463; DB 4; Length 168;
Best Local Similarity 99.0%; Pred. No. 5.2e-38;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||
Db 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||

; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

Query Match      97.3%; Score 464; DB 4; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.3e-38;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||
Db 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||
Db 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||

RESULT 4
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match      97.1%; Score 463; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 5.1e-36;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 60
    |||||||
Db 6094 LKEIDSESDYVKEGERAPLQSELDKAKQAKLSKLEESDKXKDLDAEIAKLEKXVEDFK 6153
    |||||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 100
    |||||||
Db 6154 NSDGEQAGQYLAAAEEDLIAKAKLEKAEADLKKAVIDEPE 6193
    |||||||

RESULT 5
US-08-710-749-5
; Sequence 5, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
```

;; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
;; TITLE OF INVENTION: PROTEINS  
;; NUMBER OF SEQUENCES: 28  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Curtis, Morris & Safford  
;; STREET: 530 Fifth Avenue  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/710,749  
;; FILING DATE: 20-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 101 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
US-08-710-749-5

Query Match 96.5%; Score 460.5; DB 2; Length 101;  
Best Local Similarity 99.0%; Pred. No. 5.1e-38;  
Matches 100; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
QY 1 LKEIDSESDYVKEGERAPLOSEL-DAKQAKSLKEEXSDKXDELDAEIAKLEKVDVF 59  
DB 1 LKEIDSESDYVKEGERAPLOSELDAKQAKSLKEEXSDKXDELDAEIAKLEKVDVF 60  
  
QY 60 KNSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 100  
DB 61 KNSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 101

RESULT 6  
US-09-147-875A-5  
; Sequence 5, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-5

Query Match 95.6%; Score 456; DB 4; Length 100;  
Best Local Similarity 94.0%; Pred. No. 1.4e-37;  
Matches 94; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 LKEIDSESDYVKEGERAPLOSELDAKQAKSLKEEXSDKXDELDAEIAKLEKVDVFK 60  
DB 1 LKEIDSESDYVKEGERAPLOSELDAKQAKSLKEEXSDKXDELDAEIAKLEKVDVFK 60

QY 61 NSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 100  
DB 61 NSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 100  
  
RESULT 7  
US-08-710-749-3  
; Sequence 3, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-3

Query Match 95.1%; Score 453.5; DB 2; Length 101;  
Best Local Similarity 95.0%; Pred. No. 2.5e-37;  
Matches 96; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
  
QY 1 LKEIDSESDYVKEGERAPLOSEL-DAKQAKSLKEEXSDKXDELDAEIAKLEKVDVF 59  
DB 1 LKEIDSESDYVKEGERAPLOSELDAKQAKSLKEEXSDKXDELDAEIAKLEKVDVF 60  
  
QY 60 KNSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 100  
DB 61 KNSDGEQAGQYLAAREEDLIAKXLEKAEADLKKAVIDEPE 101

RESULT 8  
US-08-529-055-50  
; Sequence 50, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yotter, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions

```

RESULT 9
US-08-710-749-4
; Sequence 4, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

CITY: NEW YORK  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/529,055  
 FILING DATE: 15-SEP-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454312-2400  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 INFORMATION FOR SEQ ID NO: 64:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 194 amino acids



```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64

Query Match      91.8%; Score 438; DB 4; Length 194;
Best Local Similarity 91.9%; Pred. No. 1.7e-35;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||
Db 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||

QY 61 NSDGEQAGQYLAAREEDLIAXKALEKAEADLKKAADLKKAVDEP 99
   |||||
Db 61 XSDGEQAGQYLAAREEDLIAXKALEKAEADLKKAADLKKAVNEP 99
   |||||

RESULT 11
US-09-147-875A-1
; Sequence 1, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-1

Query Match      88.3%; Score 421; DB 4; Length 98;
Best Local Similarity 92.0%; Pred. No. 3.6e-34;
Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||
Db 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||

QY 61 NSDGEQAGQYLAAREEDLIAXKALEKAEADLKKAADLKKAVDEPE 100
   |||||
Db 61 NSDGEQA-QYLAAREEDL-AKXAELEKTEADLKKAADLKKAVHEPE 98
   |||||

RESULT 12
US-09-583-110-4871
; Sequence 4871, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4871
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4871
```

```

Query Match      88.3%; Score 421; DB 4; Length 550;
Best Local Similarity 88.0%; Pred. No. 2.6e-33;
Matches 88; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||
Db 144 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 203
   |||||

QY 61 NSDGEQAGQYLAAREEDLIAXKALEKAEADLKKAADLKKAVDEPE 100
   |||||
Db 204 NSNGEQAEQYRAAREEDLAAKQAELEKTEADLKKAADLKKAVNEPE 243
   |||||

RESULT 13
US-09-107-433-3858
; Sequence 3858, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match      88.3%; Score 421; DB 4; Length 550;
Best Local Similarity 88.0%; Pred. No. 2.6e-33;
Matches 88; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 60
   |||||
Db 144 LKEIDSDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKQVEDPK 203
   |||||
```

SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-6  
Perfect score: 477  
Sequence: 1 LKEDSESDYVKEGERAP.....KKAXLEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	471	98.7	100	15	US-10-674-755-6
2	464	97.3	100	15	US-10-674-755-4
3	463	97.1	168	15	US-10-299-636-70
4	456	95.6	100	15	US-10-674-755-5
5	449	94.1	183	15	US-10-299-636-65
6	438	91.8	194	15	US-10-299-636-79
7	421	88.3	98	15	US-10-674-755-1
8	419	87.8	100	15	US-10-674-755-2
9	409	85.7	100	15	US-10-674-755-3
10	396	83.0	166	15	US-10-299-636-63
11	391	82.0	100	15	US-10-674-755-7
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 70, Appli
					Sequence 5, Appli
					Sequence 65, Appli
					Sequence 79, Appli
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 63, Appli
					Sequence 7, Appli

12	375.5	78.7	185	15	US-10-299-636-61	Sequence 61, Appli
13	373.5	78.3	101	15	US-10-674-755-9	Sequence 9, Appli
14	367	76.9	100	15	US-10-674-755-8	Sequence 8, Appli
15	321.5	67.4	99	15	US-10-674-755-16	Sequence 16, Appli
16	317.5	66.6	99	15	US-10-674-755-11	Sequence 11, Appli
17	317.5	66.6	204	15	US-10-299-636-66	Sequence 66, Appli
18	313.5	65.7	170	15	US-10-299-636-57	Sequence 75, Appli
19	313.5	65.7	181	15	US-10-299-636-57	Sequence 57, Appli
20	313.5	65.7	643	15	US-10-299-636-95	Sequence 95, Appli
21	313.5	65.7	670	9	US-09-748-875-63	Sequence 63, Appli
22	313.5	65.7	670	10	US-09-298-523B-63	Sequence 63, Appli
23	313.5	65.7	690	9	US-09-748-875-61	Sequence 61, Appli
24	313.5	65.7	690	10	US-09-298-523B-61	Sequence 61, Appli
25	313.5	65.7	691	9	US-09-748-875-1	Sequence 1, Appli
26	313.5	65.7	691	10	US-09-298-523B-1	Sequence 1, Appli
27	313.5	65.7	701	9	US-09-748-875-62	Sequence 62, Appli
28	313.5	65.7	701	10	US-09-298-523B-62	Sequence 62, Appli
29	313.5	65.7	707	9	US-09-748-875-2	Sequence 2, Appli
30	313.5	65.7	707	10	US-09-298-523B-2	Sequence 2, Appli
31	313.5	65.7	711	9	US-09-748-875-3	Sequence 3, Appli
32	313.5	65.7	711	10	US-09-298-523B-3	Sequence 3, Appli
33	313.5	65.7	739	17	US-10-732-923-3294	Sequence 3294, Ap
34	313.5	65.7	929	9	US-09-748-875-60	Sequence 60, Appli
35	313.5	65.7	929	10	US-09-298-523B-60	Sequence 60, Appli
36	313.5	65.7	929	15	US-10-299-636-94	Sequence 94, Appli
37	310.5	65.1	188	15	US-10-299-636-74	Sequence 74, Appli
38	310	65.0	100	15	US-10-674-755-10	Sequence 10, Appli
39	309	64.8	100	15	US-10-674-755-12	Sequence 12, Appli
40	301.5	63.2	141	14	US-10-254-995-2	Sequence 2, Appli
41	301.5	63.2	198	15	US-10-299-636-76	Sequence 76, Appli
42	301.5	63.2	354	15	US-10-299-636-105	Sequence 105, App
43	301.5	63.2	588	15	US-10-299-636-96	Sequence 96, Appli
44	301.5	63.2	589	9	US-09-748-875-14	Sequence 14, Appli
45	301.5	63.2	589	10	US-09-298-523B-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1  
US-10-674-755-6  
; Sequence 6, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-6

Query Match	98.7%	Score 471;	DB 15;	Length 100;
Best Local Similarity	100.0%;	Pred. No. 4.4e-33;		
Matches 100;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	LKEDSESDYVKEGERAPLQSELDKQAKLSEKXSDKXDELDAIAKLEKDVDFK	60	
Db	1	LKEDSESDYVKEGERAPLQSELDKQAKLSEKXSDKXDELDAIAKLEKDVDFK	60	
Qy	61	NSDGEQAGYLAABEDLIAKAKLEKAEADLKKAVIDEPE	100	

```
Db      61 NSDGEQAGQYLAABEDLIAKKAKLEKAEADLKKAVDEPE 100

RESULT 2
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match      97.3%; Score 464; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 1.8e-32;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Qy      61 NSDGEQAGQYLAABEDLIAKKAKLEKAEADLKKAVDEPE 100
Db      61 NSDGEQAGQYLAABEDLIAKKAELEKAEADLKKAVDEPE 100

RESULT 3
US-10-299-636-70
; Sequence 70, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa at position 38 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa at position 42 is unknown
; FEATURE:
; NAME/KEY: UNSURE
```

```
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70

Query Match      97.1%; Score 463; DB 15; Length 168;
Best Local Similarity 99.0%; Pred. No. 3.8e-32;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Qy      61 NSDGEQAGQYLAABEDLIAKKAKLEKAEADLKKAVDEPE 100
Db      61 NSDGEQAGQYLAABEDLIAKKAEADLKKAVDEPE 100

RESULT 4
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match      95.6%; Score 456; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 8.9e-32;
Matches 94; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db      1 LKEIDESDSDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDFK 60
Qy      61 NSDGEQAGQYLAABEDLIAKKAKLEKAEADLKKAVDEPE 100
Db      61 NSDGEQAGQYLAABEDLIAKKAELEQTEADLKKAVHEPE 100

RESULT 5
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match          94.18; Score 449; DB 15; Length 183;
Best Local Similarity 93.0%; Pred. No. 6.5e-31;
Matches 93; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVDEPE 100
   |||||
Db 61 NSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVDEPE 100
   |||||

RESULT 6
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match          91.8%; Score 438; DB 15; Length 194;
Best Local Similarity 91.9%; Pred. No. 6.1e-30;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVDEP 99
   |||||
Db 61 XSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVNEP 99
   |||||

RESULT 7
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
```

```
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match          88.3%; Score 421; DB 15; Length 98;
Best Local Similarity 92.0%; Pred. No. 8.3e-29;
Matches 92; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGLRAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVDEPE 100
   |||||
Db 61 NSDGEQA-QYLAAREEDL-AKAELEKTEADLKKAVHEPE 98
   |||||

RESULT 8
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match          87.8%; Score 419; DB 15; Length 100;
Best Local Similarity 88.0%; Pred. No. 1.3e-28;
Matches 88; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDESDSYVKEGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||
Db 1 LKEIDESDSYVKEGLRAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVEDFK 60
   |||||

QY 61 NSDGEQAGYLAAREEDLIAXKALEKAEADLKKAVDEPE 100
   |||||
Db 61 NSNGEQAEQYRAAREEDLAAKQAELEKTEADLKKAVHEPE 100
   |||||

RESULT 9
US-10-674-755-3
; Sequence 3, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
```

```
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      85.7%; Score 409; DB 15; Length 100;
Best Local Similarity 86.0%; Pred. No. 9e-28;
Matches 86; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLQSELDAAKQAKLSKLEESDKXDELDAEIAKLEKDVDFK 60
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAAKQAKLSKLEESDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQAGQYLAAREEDLIAKXAEADLKXAVDEPE 100
Db 61 NSNGEAEQYRAAGEDIAAKQAELEKTEADLKXAVHEPE 100

RESULT 10
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63

Query Match      83.0%; Score 396; DB 15; Length 166;
Best Local Similarity 82.0%; Pred. No. 2e-26;
Matches 82; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLQSELDAAKQAKLSKLEESDKXDELDAEIAKLEKDVDFK 60
Db 1 LKEIDESDSEYVKEGLRAPLQSKLDAAKAKLSKLEESDKIDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQAGQYLAAREEDLIAKXAEADLKXAVDEPE 100
Db 61 NSDGEQAGQYLVAAEKDLDAEAEALGNTGADLKXAVDEPE 100

RESULT 11
US-10-674-755-7
; Sequence 7, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)...(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-7

Query Match      82.0%; Score 391; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 3.2e-26;
Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGERAPLQSELDAAKQAKLSKLEESDKXDELDAEIAKLEKDVDFK 60
Db 1 LKEIDESDSEYKEGLRAPLQSKLDAAKAKLSKLEESDKXDELDAEIAKLEKDVDFK 60

Qy 61 NSDGEQAGQYLAAREEDLIAKXAEADLKXAVDEPE 100
Db 61 NSDGEQAGQYLVAAEKDLDAEAEALGNTGADLKXAVDEPE 100

RESULT 12
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-61

Query Match      78.7%; Score 375.5; DB 15; Length 185;
Best Local Similarity 79.2%; Pred. No. 1.3e-24;
Matches 80; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYVKEGERAPLQSELDAAKQAKLSKLEESDKXDELDAEIAK-LEKQVDF 59
Db 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAKLNKKQVDF 60

Qy 60 KNSDGEQAGQYLAAREEDLIAKXAEADLKXAVDEPE 100
Db 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKXAVNEPE 101

RESULT 13
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
```

```
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-9

Query Match      78.3%; Score 373.5; DB 15; Length 101;
Best Local Similarity 79.2%; Pred. No. 1e-24;
Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEEXSDKXDELDAETAKLEKDVDF 59
Db 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLKLSELSDKIDELDAETAKNLKDVDF 60

QY 60 KNSDGEQAGQYLAABEDLIAKXAEADLKKAVDEPE 100
Db 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101

RESULT 14
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-8

Query Match      76.9%; Score 367; DB 15; Length 100;
Best Local Similarity 79.0%; Pred. No. 3.6e-24;
Matches 79; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEEXSDKXDELDAETAKLEKDVDF 60
Db 1 LKGIIDSDSDYVKEGLRAPLOSELDAKTKLSTLELSDKIDELDAETPKLEKNVDFK 60

QY 61 NSDGEQAGQYLAABEDLIAKXAEADLKKAVDEPE 100
Db 61 LTDAEQTEQYLAABEKDLADKAELEKTEADLKKAVHEPE 100

RESULT 15
US-10-674-755-16
; Sequence 16, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-16
```

```
Query Match      67.4%; Score 321.5; DB 15; Length 99;
Best Local Similarity 69.0%; Pred. No. 2.8e-20;
Matches 69; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGERAPLOSELDAKQAKLSKLEEXSDKXDELDAETAKLEKDVDF 60
Db 1 LKEIDSDSDYVKEGERAPLOSKLDTTKAKUSKLEELSDKIDELDAETAKLEVLKDAE 60

QY 61 NSDGEQAGQYLAABEDLIAKXAEADLKKAVDEPE 100
Db 61 GNNVVEA-YFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
```

Search completed: June 18, 2005, 18:00:22  
Job time : 62.963 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-6  
Perfect score: 477  
Sequence: 1 LKEIDSESDYKGERAP.....KKAXLEKAEADLKKAVIDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: Pir1:.\*  
2: Pir2:.\*  
3: Pir3:.\*  
4: Pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	301.5	63.2	619	A97887	surface protein ps
2	301.5	63.2	619	A41971	surface protein ps
3	114	23.9	886	H69378	conserved hypothet
4	113.5	23.8	744	F93013	pneumococcal surfa
5	105.5	22.1	896	S43074	epidermal growth f
6	105	22.0	2139	T18296	myosin heavy chain
7	101.5	21.3	764	T05409	hypothetical prote
8	101	21.2	1190	E84193	chromosome segrega
9	100	21.0	1110	I51116	NF-180 - sea lamp
10	100	21.0	1269	F84730	probable myosin he
11	99	20.8	501	A44643	M protein precursor
12	99	20.8	1169	A64505	P115 homolog - Met
13	98	20.5	3488	T34418	hypothetical prote
14	97.5	20.4	897	A54696	EGF receptor subst
15	97	20.3	564	A60115	M protein precursor
16	96.5	20.2	1879	S03166	myosin heavy chain
17	96	20.1	1093	S63717	hypothetical prote
18	95.5	20.0	1927	A59236	embryonic muscle m
19	94.5	19.8	1053	A41642	dynactin - chicken
20	94	19.7	387	S57834	fcrA protein precu
21	93.5	19.6	407	EDBEQ3	immediate-early pr
22	93.5	19.6	629	T44607	hypothetical prote
23	93.5	19.6	1974	T30010	Hyp4 protein - Str
24	93	19.5	388	A46173	plasmaogen-bindin
25	93	19.5	388	A49545	M protein precursor
26	93	19.5	436	S30284	plasmaogen-bindin
27	93	19.5	454	S43556	M protein precursor
28	93	19.5	483	A26297	M6 protein - Strept
29	93	19.5	484	S46489	M1 protein precurs

30	93	19.5	532	2	S54871	M protein - Strept
31	93	19.5	539	2	A28549	M24 protein precu
32	93	19.5	688	2	A49318	protein kinase (EC
33	93	19.5	798	2	T33022	hypothetical prote
34	92	19.3	408	2	S30283	protein M precursor
35	92	19.3	710	2	AE1956	hypothetical prote
36	92	19.3	1006	2	C70445	Arpase subunit of
37	92	19.3	1690	2	T13030	microtubule bindin
38	91.5	19.2	518	2	G84488	En/Spm-like transp
39	91.5	19.2	911	2	S51441	hypothetical prote
40	91	19.1	161	2	S48396	tropomyosin TPM2
41	91	19.1	166	2	S73342	hypothetical prote
42	91	19.1	405	2	A33939	Fc gamma (IgG) rec
43	91	19.1	472	2	S43554	plasmaingen-bindin
44	91	19.1	522	2	G02533	occludin - human
45	91	19.1	1156	2	B70356	chromosome assembl

ALIGNMENTS

RESULT 1

A97887  
surface protein psps precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:9  
C:Genetics:  
A:Gene: psps

Query Match 63.2%; Score 301.5; DB 2; Length 619;  
Best Local Similarity 62.4%; Pred. No. 2.4e-14;  
Matches 68; Conservative 9; Mismatches 13; Indels 19; Gaps 2;

QY	1	LKEIDSESDYKGERAPLOSELDAKQAKLSKEEXSDKXDELDAETAKLE-----	53
DB	223	LKEIDSESDYKGERAPLOSELDAKQAKLSKEEXSDKXDELDAETAKLE-----	282
QY	54	--KVEDFKNSDGEQAGVLAAREEDLAKAKLEKAEADLKKAVIDEPE 100	
DB	283	ENNVEDY-----FKGLEKTIAAKAELEKTEADLKKAVIDEPE 321	

RESULT 2

A41971  
surface protein psps precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of Psps, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:9153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991

A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability  
A;Reference number: A60282; MUID:91169598; PMID:2004810  
A;Accession: A60282  
A;Molecule type: protein  
A;Residues: 32-76 <TAL>  
A;Experimental source: strain JV2008  
C;Genetics:  
A;Gene: pspA  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-619/Product: surface protein pspA #status predicted <MAT>  
F:411-430/Domain: cpl repeat homology <CP01>  
F:431-450/Domain: cpl repeat homology <CP02>  
F:451-470/Domain: cpl repeat homology <CP03>  
F:471-490/Domain: cpl repeat homology <CP04>  
F:491-510/Domain: cpl repeat homology <CP05>  
F:511-530/Domain: cpl repeat homology <CP06>  
F:531-550/Domain: cpl repeat homology <CP07>  
F:551-570/Domain: cpl repeat homology <CP08>  
F:571-591/Domain: cpl repeat homology <CP09>  
F:592-611/Domain: cpl repeat homology <CP10>

Query Match 63.2%; Score 301.5; DB 2; Length 619;  
Best Local Similarity 62.4%; Pred. No. 2.4e-14;  
Matches 68; Conservative 9; Mismatches 13; Indels 19; Gaps 2;

Qy 1 LKEIDSESDYVKEGERAPLOSELDAKQAKLSKLEKSDKXDELDAEIAKLE-----53

Db 223 LKEIDSESDYVKEGERAPLOSELDAKQAKLSKLEKSDKXDELDAEIAKLE-----53

Qy 54 --KQVEPFKNSDQAGQYLAAREEDLAKKXLEKAEADLKKAVDPE 100

Db 283 ENNVEDY-----FKGLEKTTAAKKAELTEADLKKAVNSPE 321

RESULT 3

H69378

conserved hypothetical protein AF1032 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69378

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artisch, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:92689355; PIDN:AAB9021

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 23.9%; Score 114; DB 2; Length 886;

Best Local Similarity 27.7%; Pred. No. 0.61;

Matches 38; Conservative 22; Mismatches 35; Indels 42; Gaps 3;

Qy 1 LKEIDSESDYVKEGERAPLOSELDAKQAKLSKLEKSDKXDELDAEIAKLEKDVDFK 60

Db 296 LSEINQALRDVKEKRGD---LTREAAGIQALKKAEDNSKLEITRKIELELERFER 352

Qy 61 NSDG-----EOAGQYLAAREEDLAKKXLEKAEADLKKAVDPE 81

Db 353 KSHRLLETLKPMRDMQIKAKLEKNTLPDKVEMWYLLSKAKEEKEITEKLKLIK 412

Qy 82 KAXLEKAEADLKKAVDE 98

Db 413 KSSLKTRGAQLKKAAVEE 429

RESULT 4

F95013

pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C;Accession: F95013

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,  
nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95013

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-744 <KUR>

A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:914971584; GSPDB:G

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP0117

Query Match 23.8%; Score 113.5; DB 2; Length 744;

Best Local Similarity 29.8%; Pred. No. 0.56;

Matches 37; Conservative 20; Mismatches 26; Indels 41; Gaps 5;

Qy 10 EDYVKEGERAPLOSELDAKQAKLSKLE-----EXSDKX 42

Db 248 EAKLKGE-----AELNAKQALAKKQTELEKLLSDLPKGTQDELDAEAEALDKKA 302

Qy 43 DELDAEIAKLEKVEDPK---NSDQAGQYLAAREEDLAKKXLEKAEADLKKAVD- 97

Db 303 DELQNKVADLEKELSLNLEILLGGADPEDD---TALQNKLAAKKAEALAKKQTELEKLLDS 359

Qy 98 -EPE 100

Db 360 LDPE 363

RESULT 5

S43074

epidermal growth factor receptor substrate - human

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S43074; I38525

R;Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.

Oncogene 9, 1039-1045, 1994

A;Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AP-4,

A;Reference number: S43074; MUID:94181254; PMID:8134107

A;Accession: S43074

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-896 <BER>

A;Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:9470034; PIDN:CAA82305.1; PID:94700

R;Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner,

Oncogene 9, 1591-1597, 1994

A;Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in evo

A;Reference number: I38525; MUID:94239734; PMID:8183552

A;Accession: I38525

A;Status: preliminary; translated from GE/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-821, 'M', 823-896 <RES>

A;Cross-references: EMBL:U07707; NID:9466259; PIDN:AAA52101.1; PID:9466260

C;Genetics:

A;Gene: GDB:EPF15; AF-1P; MLLT5

A;Cross-references: GDB:360337; OMIM:600051

A;Map position: 1p32-1p32

Query Match 22.1%; Score 105.5; DB 2; Length 896;

Best Local Similarity 27.5%; Pred. No. 2.5;

Matches 28; Conservative 28; Mismatches 39; Indels 7; Gaps 3;

Qy 3 EIDESDSEYVKE--GERAPLOSELDAKQAKLSKLE-----EXSDKXDELDAEIAKLEKDV 56

```
Db 353 EQDLKKEKEDTIKQRTSEVDQLQVQRENTNNLQKLAQKQKQVQELLDDELQKQAEQL 412
QY 57 EDFKNSDGEQAGQYLAABEDLIAKXLEKAEADLKKADE 98
Db 413 KEVRKKCAEEA-QLISSLKAEITQSBSQISTYEELAKAREE 453

RESULT 6
T18296
myosin heavy chain - Entamoeba histolytica
C:Species: Entamoeba histolytica
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18296
R:Guillen, N.
submitted to the EMBL Data Library, February 1997
A:Reference number: Z18865
A:Accession: T18296
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2139 <GUI>
A:Cross-references: UNIPROT:Q07569; EMBL:L03534; NID:g1850912; PID:g1850913; PIDN:AAB480
C:Genetics:
A:Gene: mhca
C:Superfamily: myosin heavy chain; myosin motor domain homology
F;91-780/Domain: myosin motor domain homology <MMO>

Query Match 22.0%; Score 105; DB 2; Length 2139;
Best Local Similarity 25.2%; Pred. No. 6.2;
Matches 34; Conservative 26; Mismatches 35; Indels 40; Gaps 4;

QY 2 KEIDSDSDYKVEGERAPLOSELDAKQAKL-----SKLEEXSDKXDELDAETA 50
Db 1339 KSVVESKNKD--SENEKAALSEIDQANEKLNIOADLRKATADLQEAANEKKAEEVAQRD 1396
QY 51 KLEKO-----VEDF-----KNSDGEQAGQYLAABEDLIAKKA 83
Db 1397 KLVAADNKKMTLEETKARDEENTYKVENYKVLKREKADLEAENLNDIEKKDRNRK 1456
QY 84 XLEKAEADLKKADE 98
Db 1457 QVKLEGLKETYDK 1471

RESULT 7
T05409
hypothetical protein F10M6.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05409
R:Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15414
A:Accession: T05409
A:Molecule type: DNA
A:Residues: 1-764 <BEV>
A:Experimental source: UNIPROT:O49371; EMBL:AL021811
A:Map position: 4
A:Note: F10M6.170

Query Match 21.3%; Score 101.5; DB 2; Length 764;
Best Local Similarity 30.7%; Pred. No. 4.1;
Matches 35; Conservative 21; Mismatches 41; Indels 17; Gaps 4;

QY 2 KEIDSDSDYKVEGERAPLOSELDAKQAKLKL-EEXSDKXDELDAETIAKLE----- 53
Db 163 REIEELKHKLREDBERAAQLQSLTLKEBELEKMRQETANRSKVSMAISBFESKQLLS 222
QY 54 KDVEFKNSDGE--QAGQYLAABEDLIAKX-----LEKAEADLKKADE 98
Db 223 KANENVKQGEIYALQRALEEKBELETSKATKLEQKLEFTEANLKKQTEE 276
```

## RESULT 8

```
EMBL193
Chromosome segregation [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84193
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1190 <STO>
A:Cross-references: UNIPROT:Q9HS95; GB:AE004437; NID:g10579965; PIDN:AAG18913.1; GSPDB:
C:Genetics:
A:Gene: smc1
```

## Query Match

21.2%; Score 101; DB 2; Length 1190;

Best Local Similarity 34.3%; Pred. No. 6.7;

Matches 37; Conservative 20; Mismatches 35; Indels 16; Gaps 5;

QY 3 EIDESDSDYKVEGERA---PLQSELDKQAKLXKLEEXSDKXDELDAETIAKLEKDVDF 59

Db 858 QADVADASR-KADQARIEALNGDIEAKQALAEKAA---VEDLEAELADLKRDREER 913

QY 60 KN--SDGQAGQYLAABED-----LIAKKXLEKAEADLKKADE 98

Db 914 KADLSEADARDEQAQAAVEDARHRLERLQAAQTLSSEVAELDDAVGE 961

## RESULT 9

```
IF1116
NF-180 - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: IF1116
R:Jacobs, A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A:Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation
A:Reference number: IF1116; MUID:95287814; PMID:7770000
A:Accession: IF1116
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1110 <JAC>
A:Cross-references: UNIPROT:O91255; EMBL:U19361; NID:g632548; PIDN:AAA80106.1; PID:g632
C:Superfamily: neurofilament triplet H protein
```

## Query Match

21.0%; Score 100; DB 2; Length 1110;

Best Local Similarity 25.3%; Pred. No. 7.4;

Matches 25; Conservative 34; Mismatches 34; Indels 6; Gaps 3;

QY 2 KEIDSD--SEDYKVEGERAPLOSELDAKQAKLXKLEEXSDKXDELDAEI---AKLEKDV 56

Db 568 BEAEAEVEBEATEKAEAAKAEVEEEAEAEVEEEAEVEEAEVEEAEVEEAEVEEAEV 627

QY 57 EDFKNSDGEQAGQYLAABEDLIAKXLEKAEADLKKA 95

Db 628 EE-EGEAAEEAEAEAEAEVEEVTSKAKTQAEVEEEEA 665

## RESULT 10

```
F84730
probable myosin heavy chain [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84730
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
```



```
A::Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 29-89 <WHW>
A:Cross-references: EMBL:U11937
C:Genetics:
A:Gene: emm12
C:Superfamily: M5 protein
C:Keywords: transmembrane protein
F:1-41/Domain: signal sequence #status pro
F:42-564/Product: M protein (fragment) #sig
```

	Query Match	20.3%	Score 9	
	Best Local Similarity	26.1%	Pred NM	
	Matches	24;	Conservative	21; Mismatches
Qy	3	EIDSSDSEYVKEGERAPLOSLDLAKQKQAEI	:	:
Dd	403	ELDKVBEKQISDSARQGRLRRDLDSRIR	:	:
Qy	63	DGEAGOVLAIAAESLDLIAKKXLEKAEEI	:	:
Dd	463	TEKSXAELQAQLKFAVKALSKQLAKQA	:	:

Search completed: June 18, 2005, 17:03:53  
Job time : 13.113 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-6  
Perfect score: 477  
Sequence: 1 LKEIDSDSEYVKEGERAP.....KXAXLEKAEADLKXAVDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	97.1	394	2 Q9LAY6	Q9lay6 streptococc
2	463	97.1	395	2 Q9LAZ1	Q9laz1 streptococc
3	450	94.3	406	2 Q9LAZ0	Q9laz0 streptococc
4	443	92.9	340	2 Q8KOK5	Q8kok5 streptococc
5	426	89.3	225	2 Q9L591	Q9l591 streptococc
6	421	88.3	222	2 Q9L577	Q9l577 streptococc
7	421	88.3	262	2 Q9L576	Q9l576 streptococc
8	421	88.3	415	2 Q9LAY7	Q9lay7 streptococc
9	418	87.6	194	2 Q9L5B5	Q9l5b5 streptococc
10	418	87.6	218	2 Q6UBB2	Q6ueb2 streptococc
11	418	87.6	233	2 Q9L568	Q9l568 streptococc
12	418	87.6	236	2 Q9L569	Q9l569 streptococc
13	418	87.6	243	2 Q9L564	Q9l564 streptococc
14	418	87.6	243	2 Q9L567	Q9l567 streptococc
15	418	87.6	244	2 Q9L565	Q9l565 streptococc
16	418	87.6	247	2 Q9L566	Q9l566 streptococc
17	418	87.6	249	2 Q9L570	Q9l570 streptococc
18	418	87.6	254	2 Q9L563	Q9l563 streptococc
19	418	87.6	401	2 Q9LAZ2	Q9laz2 streptococc
20	414	86.8	255	2 Q9L581	Q9l581 streptococc
21	414	86.8	255	2 Q9L5B6	Q9l5b6 streptococc
22	414	86.8	416	2 Q9LAY8	Q9lay8 streptococc
23	411	86.2	246	2 Q9L578	Q9l578 streptococc
24	394	82.6	393	2 Q9LAZ3	Q9laz3 streptococc
25	383	80.3	207	2 Q8GNS9	Q8gns9 streptococc
26	377	79.0	237	2 Q9L592	Q9l592 streptococc
27	377	79.0	395	2 Q9LAY9	Q9lay9 streptococc
28	317	66.6	417	2 Q9LAY3	Q9lay3 streptococc
29	313.5	65.7	739	2 Q9RQT4	Q9rqt4 streptococc
30	313.5	65.7	820	2 Q9RQT1	Q9rqt1 streptococc
31	313.5	65.7	929	2 Q9KK19	Q9kk19 streptococc

32 313.5 65.7 929 2 Q9ZAY5 streptococc  
33 301.5 63.2 619 2 Q54972 streptococc  
34 301.5 63.2 619 2 Q8DRI0 streptococc  
35 298.5 62.6 249 2 Q9L575 streptococc  
36 297.5 62.4 415 2 Q9LAY1 streptococc  
37 297.5 62.4 437 2 Q9LAY4 streptococc  
38 294.5 61.7 99 2 Q8KQK4 streptococc  
39 291.5 61.1 224 2 Q8GNS8 streptococc  
40 291.5 61.1 426 2 Q9LAY5 streptococc  
41 286.5 60.1 395 2 Q9LAY2 streptococc  
42 286.5 60.1 408 2 Q9LAY0 streptococc  
43 278 58.3 869 2 Q9KK27 streptococc  
44 173.5 36.4 479 2 Q9LAX2 streptococc  
45 173.5 36.4 480 2 Q9LAX3 streptococc

## ALIGNMENTS

### RESULT 1

Q9LAY6 PRELIMINARY; PRT; 394 AA.  
AC Q9LAY6;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
OS Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L81905;  
RX MEDLINE=20448953; PubMed=10992499;  
RX DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae";  
RL Infect Immun. 68:5889-5900(2000).  
DR EMBL; AF071809; AAF27705.1; -  
DR InterPro; IPR009082; His\_kin\_homodim.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 394 394  
SQ SEQUENCE 394 AA; 42935 MW; A192F47E12C89FAA CRC64;

Query Match 97.1%; Score 463; DB 2; Length 394;  
Best Local Similarity 96.0%; Pred. No. 4.4e-24;  
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSEYVKEGERAPQSELDQAQAKLSKLEESDKYDELDAETAKLEKQVEDPK 60  
Db 213 LKEIDSDSEYVKEGFRAPQSELDQAQAKLSKLEESDKYDELDAETAKLEKQVEDPK 272

QY 61 NSDQGAQGYLAAREEDLIKAKLEKAEADLKXAVDEPE 100  
Db 273 NSDQGAQGYLAAREEDLIKAKLEKAEADLKXAVDEPE 312

### RESULT 2

Q9LAZ1 PRELIMINARY; PRT; 395 AA.  
AC Q9LAZ1;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
OS Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```

OC Streptococcus
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9739;
RX MEDLINE=20448953; PubMed=10992499;
RY DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071804; AAF27700.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42968 MW; 67FDECA41DB7F95 CRC64;

Query Match 97.1%; Score 463; DB 2; Length 395;
Best Local Similarity 96.0%; Pred. No. 4.4e-24;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db 213 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 272

Oy 61 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 100
Db 273 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 312

RESULT 3
O9LAZO PRELIMINARY; PRT; 406 AA.
AC O9LAZO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DLB6A;
RX MEDLINE=20448953; PubMed=10992499;
RY DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 94.3%; Score 450; DB 2; Length 406;
Best Local Similarity 93.0%; Pred. No. 3.5e-23;
Matches 93; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db 213 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 272

Oy 61 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 100
Db 273 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 312

```

```

RESULT 4
O8KQK5 PRELIMINARY; PRT; 340 AA.
AC O8KQK5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
RY DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
immunization with DNA vaccines against Streptococcus pneumoniae
expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 340
SQ SEQUENCE 340 AA; 38023 MW; EE07ECF00B1FBD57 CRC64;

Query Match 92.9%; Score 443; DB 2; Length 340;
Best Local Similarity 91.0%; Pred. No. 8.9e-23;
Matches 91; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Oy 1 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 60
Db 197 LKEIDESDYVKEGERAPLQSELDKQAKLSKLEEXSDKXDELDAEIAKLEKDVDFK 256

Oy 61 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 100
Db 257 NSDGEAQGYLAAAEEDLIAKAKLEKAEADLKXAVDEPE 296

RESULT 5
O9L591 PRELIMINARY; PRT; 225 AA.
AC O9L591;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RY Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RY Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```





```
||||| 289 NSNGEQAEQYRAAAEDLAAQAELEKTEADLKKAVDEPE 328
Db
RESULT 9
Q9L5B5
ID Q9L5B5 PRELIMINARY; PRT; 194 AA.
AC Q9L5B5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253407; AAF67355.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 194
SQ SEQUENCE 194 AA; 21116 MW; E68189FCA2B244F8 CRC64;

Query Match 87.6%; Score 418; DB 2; Length 194;
Best Local Similarity 86.0%; Pred. No. 2.8e-21;
Matches 86; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGERAPLQSELDAAQAKLSKLEEXSDKXDELDAEIAKLEKXVDFK 60
|||||
Db 55 LKEIDSDSDYIKGLRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLEKXVDFK 114
|||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKKALEKAEADLKKAVDEPE 100
|||||
Db 115 NSDGEQAEQYLVAACKLDLDAKKAELNTEADLKKAVDEPE 154
|||||

RESULT 10
Q6UEB2
ID Q6UEB2 PRELIMINARY; PRT; 218 AA.
AC Q6UEB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URSP2;
RX PubMed=14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;
RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;
RT "Epitope mapping of a protective monoclonal antibody against
RT pneumocystis carinii with shared reactivity to Streptococcus
RT pneumoniae surface antigen PspA.";
RL Infect. Immun. 72:1548-1556(2004).
DR EMBL; AY371665; AAR20918.1; -.

DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 218
SQ SEQUENCE 218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;

Query Match 87.6%; Score 418; DB 2; Length 218;
Best Local Similarity 86.0%; Pred. No. 3.1e-21;
Matches 86; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGERAPLQSELDAAQAKLSKLEEXSDKXDELDAEIAKLEKXVDFK 60
|||||
Db 27 LKEIDSDSDYIKGLRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLEKXVDFK 86
|||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKKALEKAEADLKKAVDEPE 100
|||||
Db 87 NSDGEQAEQYLVAACKLDLDAKKAELNTEADLKKAVDEPE 126
|||||

RESULT 11
Q9L568
ID Q9L568 PRELIMINARY; PRT; 233 AA.
AC Q9L568;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255902; AAF70092.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 233
SQ SEQUENCE 233 AA; 24514 MW; D5C494019C45BFE2 CRC64;

Query Match 87.6%; Score 418; DB 2; Length 233;
Best Local Similarity 86.0%; Pred. No. 3.2e-21;
Matches 86; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGERAPLQSELDAAQAKLSKLEEXSDKXDELDAEIAKLEKXVDFK 60
|||||
Db 28 LKEIDSDSDYIKGLRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLEKXVDFK 87
|||||

Qy 61 NSDGEQAGQYLAAAEEDLIAKKALEKAEADLKKAVDEPE 100
|||||
Db 88 NSDGEQAEQYLVAACKLDLDAKKAELNTEADLKKAVDEPE 127
|||||

RESULT 12
Q9L569
ID Q9L569 PRELIMINARY; PRT; 236 AA.
AC Q9L569;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
```



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=183;  
 RX MEDLINE=20472698; PubMed=11015380;  
 RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
 RT "Pneumococcal psppa sequence types of prevalent multiresistant  
 RT pneumococcal strains in the United States and of internationally  
 RT disseminated clones."; 38:3663-3669(2000).  
 RL J. Clin. Microbiol. 38:3663-3669(2000).  
 RN [2]

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=183;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF255905; AAF70095.1; -;  
 DR InterPro; IPR009082; His\_kin\_homodim.  
 FT NON\_TER 1  
 FT NON\_TER 244 244  
 SQ SEQUENCE 244 AA; 25946 MW; F9274FFD1957DD06 CRC64;

Query Match 87.6%; Score 418; DB 2; Length 244;  
 Best Local Similarity 86.0%; Pred. NO. 3.4e-21;  
 Matches 86; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGEAPLOQSLDAKQAKLSKLEEXSDKXDELDAEIAKLEKQVDEPK 60  
 Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 51 LKEIDSDSDYIIEGLRAPLOSLDAKQAKLSKLEELSDKIDELDAEIAKLEKQVDEPK 110  
 Qy 61 NSDGEQAGQYLAAAEEDLIAKKALEKAEADLKKAVDEPE 100  
 Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 111 NSDGEQAEQYLVAAKKDLDAKKALENTADLKKAVDEPE 150

Search completed: June 18, 2005, 17:01:35  
 Job time : 60.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-7

Perfect score: 491

Sequence: 1 LKEIDESSEDIYEKGLRAP.....KEAELNTGADLKKAVIDPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	97.6	166	AAW14568	Aaw14568 Streptococ
2	479	97.6	166	ABW02602	Abw02602 Bg8743c p
3	479	97.6	8991	ABU08487	Abu08487 S. pneumo
4	394	80.2	168	ABW02609	Abw02609 L81905c p
5	390	79.4	183	AAW14570	Aaw14570 Streptococ
6	390	79.4	183	ABW02604	Abw02604 Bg9739c p
7	388	79.0	194	AAW14584	Aaw14584 Streptococ
8	388	79.0	194	ABW02618	Abw02618 Db16ac pn
9	382	77.8	550	ADK48356	Adk48356 Streptococ
10	382	77.8	550	ADR95223	Adr95223 Novel S.
11	378.5	77.1	167	AAW14575	Aaw14575 Streptococ
12	341.5	69.6	185	AAW14566	Aaw14566 Streptococ
13	341.5	69.6	185	ABW02600	Abw02600 Ac94c pne
14	297.5	60.6	170	ABW02614	Abw02614 Rct135c p
15	297.5	60.6	181	ABW02596	Abw02596 0922134c
16	297.5	60.6	865	ABU08489	Abu08489 S. pneumo
17	297.5	60.6	929	AAW14593	Aaw14593 Streptococ
18	297.5	60.6	929	AAY43384	Aay43384 S. pneumo
19	294.5	60.0	188	AAW14580	Aaw14580 Streptococ
20	294.5	60.0	188	ABW02613	Abw02613 Rct129c p
21	294.5	60.0	198	AAW14581	Aaw14581 Streptococ
22	291.5	59.4	198	ABW02615	Abw02615 Rxl1c pneu
23	291.5	59.4	315	AAW04375	Aay04375 Streptococ
24	291.5	59.4	619	AAR63437	Aar63437 Pneumococ
25	291.5	59.4	619	AAR87598	Aar87598 Pneumococ

26	291.5	59.4	619	2	AAR86911	Aar86911 Pneumococ
27	291.5	59.4	619	2	AAY41838	Aay41838 Streptococ
28	291.5	59.4	619	5	AAE18782	Aae18782 S. pneumo
29	291.5	59.4	619	6	ABU45778	Abu45778 Protein e
30	291.5	59.4	619	8	ADO52126	Ado52126 Streptococ
31	291.5	59.4	648	2	AAW70336	Aaw70336 Pneumococ
32	291.5	59.4	648	2	AAW62274	Aaw62274 Streptococ
33	291.5	59.4	648	2	AAY41837	Aay41837 Streptococ
34	291.5	59.4	648	2	AAW87879	Aaw87879 A. pneumoc
35	291.5	59.4	653	2	AAW92456	Aaw92456 S. pneumo
36	291.5	59.4	684	2	AAR73912	Aar73912 Streptococ
37	286.5	58.4	204	2	AAW14571	Aaw14571 Streptococ
38	286.5	58.4	204	7	ABW02605	Abw02605 Bf1019c p
39	286	58.2	180	2	AAW14562	Aaw14562 Streptococ
40	285.5	58.1	588	6	ABU08491	Abu08491 Coiled co
41	285.5	58.1	589	2	AAY43392	Aay43392 PspC alph
42	283.5	57.7	195	2	AAW14591	Aaw14591 Streptococ
43	283.5	57.7	195	7	ABW02625	Abw02625 Wuzc pneu
44	283.5	57.7	204	2	AAW14578	Aaw14578 Streptococ
45	283.5	57.7	204	7	ABW02612	Abw02612 Rct123c p

## ALIGNMENTS

## RESULT 1

AAW14568  
ID AAW14568 standard; protein; 166 AA.  
AC AAW14568;  
XX  
DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
DE Streptococcus pneumoniae PspA central region.  
XX  
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
OS Streptococcus pneumoniae; strain Bg8743.  
XX  
PN WO9709994-A1.  
XX  
PD 20-MAR-1997.  
XX  
PF 16-SEP-1996; 96WO-US014819.  
XX  
PR 15-SEP-1995; 95US-00529055.  
XX (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiarlo E, Yorher J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.  
XX  
DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX  
PS Example 6; Fig 13; 296pp; English.  
XX  
CC This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine

CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 166 AA;

Query Match 97.6%; Score 479; DB 2; Length 166;  
 Best Local Similarity 97.0%; Pred. No. 1.1e-41;  
 Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 60  
 Db 1 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 60

Oy 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100  
 Db 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100

RESULT 2  
 ABW02602  
 ID ABW02602 standard; protein; 166 AA.  
 XX  
 AC ABW02602;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Bg8743c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 FN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 DR WPI; 2003-862841/80.  
 XX  
 PS Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain.  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 48; 121pp; English.

CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Bg8743c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention

XX SQ Sequence 166 AA;

Query Match 97.6%; Score 479; DB 7; Length 166;  
 Best Local Similarity 97.0%; Pred. No. 1.1e-41;

Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 60  
 Db 1 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 60

Oy 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100  
 Db 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100

RESULT 3  
 ABU08487  
 ID ABU08487 standard; protein; 8991 AA.  
 XX  
 AC ABU08487;  
 XX  
 DT 24-JUN-2003 (first entry)  
 XX  
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
 XX  
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..8991  
 FT /note= "All Xaa residues within this sequence are  
 FT unknown"  
 XX  
 XX US6500613-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 16-SEP-1996; 96US-00714741.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UYAL-) UNIV ALABAMA.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 2003-361534/34.  
 XX  
 PS Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 XX  
 PS Disclosure; Col 145-188; 186pp; English.

CC The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae  
 CC PspA protein

XX SQ Sequence 8991 AA;

Query Match 97.6%; Score 479; DB 6; Length 8991;  
 Best Local Similarity 97.0%; Pred. No. 1.4e-39;  
 Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 60  
 Db 4780 LKEIDSDSEDEYKEGLRAPLQSKLDAKAKLKLDESDKXKDELDAEIAKLEKDVGFPP 4839

QY 61 NSDGEQAGYLVAAEKDLDKAEELGNTGADLKKAVIDEPE 100  
 DB 4840 NSDGEQAGYLVAAEKDLDKAEELGNTGADLKKAVIDEPE 4879

RESULT 4  
 ID AEW02609  
 AC AEW02609 standard; protein; 168 AA.

XX AEW02609;

XX 12-FEB-2004 (first entry)

DE L81905c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers  
 FT Misc-difference 1..168  
 FT /note= "Xaa = Unknown amino acid"

PN US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 55; 121pp; English.

XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is L81905c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention

XX Sequence 168 AA;

Query Match 80.2%; Score 394; DB 7; Length 168;  
 Best Local Similarity 84.0%; Pred. No. 7.2e-33;  
 Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKEIDESSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVDPF 60

DB 1 LKEIDESSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVDPF 60

QY 61 NSDGEQAGYLVAAEKDLDKAEELGNTGADLKKAVIDEPE 100

|||||

DB 61 NSDGEQAGYLVAAEBEDLIAKXLEKAEADLKKAVIDEPE 100

RESULT 5  
 AAW14570

ID AAW14570 standard; protein; 183 AA.

XX AAW14570;

XX 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain Bg9739.

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 183 AA;

Query Match 79.4%; Score 390; DB 2; Length 183;  
 Best Local Similarity 81.0%; Pred. No. 2.1e-32;  
 Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKEIDESSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVDPF 60

DB 1 LKEIDESSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVDPF 60

QY 61 NSDGEQAGYLVAAEKDLDKAEELGNTGADLKKAVIDEPE 100

|||||

DB 61 NSDGEQAGYLVAAEKDLDKAEELGNTGADLKKAVIDEPE 100

RESULT 6

ID AEW02604

XX AEW02604 standard; protein; 183 AA.

XX AC AEW02604;





FH Key Location/Qualifiers  
 FT Misc-difference 1.194 /note= "Xaa = Unknown amino acid"  
 XX  
 XX US6592876-B1.  
 XX  
 XX 15-JUL-2003.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 XX WPI; 2003-862841/80.  
 XX  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 64; 121pp; English.  
 PS  
 XX  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Dbl6ac pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 194 AA;  
 Query Match 79.0%; Score 388; DB 7; Length 194;  
 Best Local Similarity 81.8%; Pred. No. 3.6e-32;  
 Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 LKEIDSDSEDYKEGLRAPLQSKLDAAKAKLSKLDKSDKXDELDAETAKLEKVDGPP 60  
 DB 1 LKEIDSDSEDYKEGLRAPLQSKLDAAKAKLSKLDKSDKXDELDAETAKLEKVDGPP 60  
 QY 61 NSDGEQAGQYLVAEKDLDAKAEELGNTGADLKAVDEP 99  
 DB 61 XSDGEQAGQYLVAEKDLDAKAEELGNTGADLKAVNEP 99  
 RESULT 9  
 ADK48356  
 ID ADK48356 standard; protein; 550 AA.  
 AC ADK48356;  
 XX  
 XX 20-MAY-2004 (first entry)  
 DT  
 DE Streptococcus pneumoniae protein, Seq ID No 4871.  
 XX  
 KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN US6699703-B1.  
 XX  
 PD 02-MAR-2004.

XX 26-MAY-2000; 2000US-00583110.  
 PF  
 XX 02-JUL-1997; 97US-0051553P.  
 PR  
 PR 12-MAY-1998; 98US-0085131P.  
 PR  
 PR 30-JUN-1998; 98US-00107433.  
 XX  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA  
 XX  
 XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;  
 PI  
 XX WPI; 2004-212399/20.  
 DR  
 DR N-PSDB; ADK45695.  
 XX  
 XX New nucleic acid molecules and polypeptides useful for diagnosing,  
 PT preventing and treating pathological conditions resulting from bacterial  
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
 PT screening.  
 XX  
 XX Disclosure; SEQ ID NO 4871; 301pp; English.  
 PS  
 XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
 CC and polypeptides. The nucleic acids and proteins are useful for  
 CC diagnosing, preventing and treating pathological conditions resulting  
 CC from bacterial infection, such as S. pneumoniae infection. These may also  
 CC be used for drug screening procedures. The present sequence represents a  
 CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
 CC data for this patent did not appear in the printed specification but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 550 AA;  
 Query Match 77.8%; Score 382; DB 8; Length 550;  
 Best Local Similarity 80.0%; Pred. No. 5.4e-31;  
 Matches 80; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 LKEIDSDSEDYKEGLRAPLQSKLDAAKAKLSKLDKSDKXDELDAETAKLEKVDGPP 60  
 DB 144 LKEIDSDSEDYKEGLRAPLQSKLDAAKAKLSKLDKSDKXDELDAETAKLEKVDGPP 203  
 QY 61 NSDGEQAGQYLVAEKDLDAKAEELGNTGADLKAVDEP 100  
 DB 204 NSNGEAGQYRAAEEDLAAKQAELEKTEADLKAVNEP 243  
 RESULT 10  
 ADR95223  
 ID ADR95223 standard; protein; 550 AA.  
 XX  
 XX ADR95223;  
 AC  
 XX 16-DEC-2004 (first entry)  
 DT  
 DE Novel S. pneumoniae protein sequence, SEQ ID 3858.  
 XX  
 KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;  
 KW bacterial infection.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN US6800744-B1.  
 XX  
 PD 05-OCT-2004.  
 XX  
 PF 30-JUN-1998; 98US-00107433.  
 XX  
 PR 02-JUL-1997; 97US-0051553P.  
 PR  
 PR 12-MAY-1998; 98US-0085131P.  
 XX  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA  
 XX  
 XX Doucette-Stamm LA, Bush D;



PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX  
 PS This sequence shows the central portion, including the C-terminus of the  
 XX alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-Oct-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 69.6%; Score 341.5; DB 2; Length 185;  
 Best Local Similarity 74.3%; Pred. No. 2.3e-27;  
 Matches 75; Conservative 7; Mismatches 18; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAK-LEKDVGF 59  
 DB 1 LKEIDESDSEYKEGLRVPLQSELDVQKAKLLELSKLDKDELDAEIAKLNKDVDF 60  
 QY 60 PNSDGEQAGYLVAEKDLDAKEAEIGNTGADLKKAVIDPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVIDPE 101  
 RESULT 13  
 ABW02600  
 ID ABW02600 standard; protein; 185 AA.  
 AC ABW02600;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR Immunological composition for obtaining expression products used for  
 XX detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding

PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 46; 121pp; English.  
 XX  
 PS The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 69.6%; Score 341.5; DB 7; Length 185;  
 Best Local Similarity 74.3%; Pred. No. 2.3e-27;  
 Matches 75; Conservative 7; Mismatches 18; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAK-LEKDVGF 59  
 DB 1 LKEIDESDSEYKEGLRVPLQSELDVQKAKLLELSKLDKDELDAEIAKLNKDVDF 60  
 QY 60 PNSDGEQAGYLVAEKDLDAKEAEIGNTGADLKKAVIDPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVIDPE 101  
 RESULT 14  
 ABW02614  
 ID ABW02614 standard; protein; 170 AA.  
 XX  
 AC ABW02614;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Rct135c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR Immunological composition for obtaining expression products used for  
 XX detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding

CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspA) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Rct135c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 170 AA;

Query Match 60.6%; Score 297.5; DB 7; Length 170;  
Best Local Similarity 67.0%; Pred. No. 7.7e-23;  
Matches 67; Conservative 7; Mismatches 25; Indels 1; Gaps 1;  
Qy 1 LKEIDESDSEYKEGLRAPLQSKLDAKAKLKLDEKSDKXDELDAEIAKLEKVDGDFP 60  
Db 1 LKEIDESDSEYKEGLRAPLQSKLDTKAKLKLDELSKIDELDAEIAKLEVLQKDAE 60  
Qy 61 NSDGEQAGQYLVAEAKDLDAEAEELGNTGADLKKAVDEPE 100  
Db 61 GNNNVEA-YFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 15  
ABW02596  
ID ABW02596 standard; protein; 181 AA.  
XX  
AC ABW02596;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE 0922134c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PW US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
DR WPI; 2003-862841/80.  
XX  
PT Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
PS Example 6; SEQ ID NO 42; 121pp; English.  
XX  
CC The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspA) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,

CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is 0922134c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 181 AA;

Query Match 60.6%; Score 297.5; DB 7; Length 181;  
Best Local Similarity 67.0%; Pred. No. 8.3e-23;  
Matches 67; Conservative 7; Mismatches 25; Indels 1; Gaps 1;  
Qy 1 LKEIDESDSEYKEGLRAPLQSKLDAKAKLKLDEKSDKXDELDAEIAKLEKVDGDFP 60  
Db 1 LKEIDESDSEYKEGLRAPLQSKLDTKAKLKLDELSKIDELDAEIAKLEVLQKDAE 60  
Qy 61 NSDGEQAGQYLVAEAKDLDAEAEELGNTGADLKKAVDEPE 100  
Db 61 GNNNVEA-YFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

Search completed: June 18, 2005, 16:51:22  
Job time : 74.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-7  
Perfect score: 491  
Sequence: 1 LKEIDSESDYEKEGLRAP.....KEAELGNTGADLKKAVDPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	99.2	100	4	US-09-147-875A-7
2	479	97.6	166	4	US-08-529-055-48
3	479	97.6	8991	4	US-08-714-741-32
4	461.5	94.0	101	2	US-08-710-749-6
5	395	80.4	100	4	US-09-147-875A-5
6	394	80.2	168	4	US-08-529-055-55
7	391	79.6	100	4	US-09-147-875A-6
8	390	79.4	100	4	US-09-147-875A-4
9	388	79.0	183	4	US-08-529-055-50
10	388	79.0	194	4	US-08-529-055-64
11	386.5	78.7	101	2	US-08-710-749-4
12	382	77.8	550	4	US-09-583-110-4871
13	382	77.8	550	4	US-09-107-433-3858
14	380.5	77.5	101	2	US-08-710-749-5
15	380	77.4	100	4	US-09-147-875A-2
16	379.5	77.3	101	2	US-08-710-749-3
17	378	77.0	98	4	US-09-147-875A-1
18	371.5	75.7	101	2	US-08-710-749-1
19	370	75.4	100	4	US-09-147-875A-3
20	369.5	75.3	99	2	US-08-710-749-9
21	364.5	74.2	101	2	US-08-710-749-2
22	341.5	69.6	185	4	US-08-529-055-46
23	339.5	69.1	101	4	US-09-147-875A-9
24	336	68.4	100	4	US-09-147-875A-8
25	331	67.4	102	2	US-08-710-749-8
26	327.5	66.7	101	2	US-08-710-749-7
27	297.5	60.6	170	4	US-08-529-055-60

28	297.5	60.6	181	4	US-08-529-055-42	Sequence 42, Appl
29	297.5	60.6	864	4	US-08-714-741-40	Sequence 40, Appl
30	295	60.1	100	4	US-09-147-875A-12	Sequence 12, Appl
31	294.5	60.0	188	4	US-08-529-055-59	Sequence 59, Appl
32	293.5	59.8	99	2	US-08-710-749-17	Sequence 17, Appl
33	292.5	59.6	99	2	US-08-710-749-15	Sequence 15, Appl
34	291.5	59.4	99	2	US-08-710-749-11	Sequence 11, Appl
35	291.5	59.4	198	4	US-08-529-055-61	Sequence 61, Appl
36	291.5	59.4	619	1	US-08-465-746-2	Sequence 2, Appl
37	291.5	59.4	619	1	US-08-214-164-2	Sequence 2, Appl
38	291.5	59.4	619	2	US-08-467-852A-3	Sequence 3, Appl
39	291.5	59.4	619	2	US-08-246-636-2	Sequence 2, Appl
40	291.5	59.4	619	2	US-08-247-491A-3	Sequence 3, Appl
41	291.5	59.4	619	2	US-08-319-795-2	Sequence 2, Appl
42	291.5	59.4	619	2	US-08-468-985-2	Sequence 2, Appl
43	291.5	59.4	619	3	US-08-312-949-2	Sequence 2, Appl
44	291.5	59.4	648	1	US-08-072-070-2	Sequence 2, Appl
45	291.5	59.4	648	1	US-08-469-434-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-147-875A-7  
; Sequence 7, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-09-147-875A-7

Query Match	99.2%	Score 487;	DB 4;	Length 100;
Best Local Similarity	100.0%;	Pred. No. 1.9e-44;		
Matches 100;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	LKEIDSESDYEKEGLRAPLQSKLDAKAKLSKLDKSDKXDELDAETAKLEKVDGDPF	60	
Db	1	LKEIDSESDYEKEGLRAPLQSKLDAKAKLSKLDKSDKXDELDAETAKLEKVDGDPF	60	
QY	61	NSDGEQAGQYLVAEKLDLDAKAEELGNTGADLKKAVDPE	100	
Db	61	NSDGEQAGQYLVAEKLDLDAKAEELGNTGADLKKAVDPE	100	

RESULT 2  
US-08-529-055-48  
; Sequence 48, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-48

Query Match 97.6%; Score 479; DB 4; Length 166;
Best Local Similarity 97.0%; Pred. No. 2.5e-43;
Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYEKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAKLEKVDGDFP 60
Db 1 LKEIDSESDYEKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAKLEKVDGDFP 60
Qy 61 NSDGEQAGQYLVAEKKDLDAKEAEELGNTGADLKKAVDEPE 100
Db 61 NSDGEQAGQYLVAEKKDLDAKEAEELGNTGADLKKAVDEPE 100

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match 97.6%; Score 479; DB 4; Length 8991;
Best Local Similarity 97.0%; Pred. No. 3.4e-41;
Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYEKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAKLEKVDGDFP 60
Db 4780 LKEIDSESDYEKEGLRAPLQSKLDKAKKLSKLDKSDKXDELDAEIAKLEKVDGDFP 4839
Qy 61 NSDGEQAGQYLVAEKKDLDAKEAEELGNTGADLKKAVDEPE 100
Db 4840 NSDGEQAGQYLVAEKKDLDAKEAEELGNTGADLKKAVDEPE 4879

RESULT 4
US-08-710-749-6
; Sequence 6, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
```

```
;
;   TOPOLOGY: linear
;   MOLECULE TYPE: amino acid
US-08-710-749-6

Query Match      94.0%; Score 461.5; DB 2; Length 101;
Best Local Similarity 97.0%; Pred. No. 9.6e-42;
Matches 98; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYKEGRLAPLQSKL-DAKAKLSKLDKSDKXDELDAETAKLEKDVGF 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDSDSDYKEGRLAPLQSKLDDAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 60 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match      80.4%; Score 395; DB 4; Length 100;
Best Local Similarity 83.0%; Pred. No. 1.1e-34;
Matches 83; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKEGRLAPLQSKLDKAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDSDSDYKEGRLAPLQSKLSDAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-08-529-055-55
; Sequence 55, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/529,055
;   FILING DATE: 15-SEP-1995
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Frommer, William S.
;   REGISTRATION NUMBER: 25,506
;   REFERENCE/DOCKET NUMBER: 454312-2400
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 840-3333
;   TELEFAX: (212) 840-0712
;   INFORMATION FOR SEQ ID NO: 55:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 168 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   US-08-529-055-55

Query Match      80.2%; Score 394; DB 4; Length 168;
Best Local Similarity 84.0%; Pred. No. 2.6e-34;
Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKEGRLAPLQSKLDKAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDSDSDYKEGRLAPLQSKLSDAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-147-875A-6
; Sequence 6, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-09-147-875A-6

Query Match      79.6%; Score 391; DB 4; Length 100;
Best Local Similarity 84.0%; Pred. No. 2.8e-34;
Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKEGRLAPLQSKLDKAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDSDSDYKEGRLAPLQSKLSDAKAKLSKLDKSDKXDELDAETAKLEKDVGF 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSDGEQAGQYLVAAEKDLDKAEALGNTGADLKKAVIDEPE 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-147-875A-4
; Sequence 4, Application US/09147875A
```

```
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-4

Query Match      79.4%; Score 390; DB 4; Length 100;
Best Local Similarity 82.0%; Pred. No. 3.6e-34;
Matches 82; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFP 60
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFK 60

Qy 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 100
Db 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 100

RESULT 9
US-08-529-055-50
; Sequence 50, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

Query Match      79.4%; Score 388; DB 4; Length 194;
Best Local Similarity 81.8%; Pred. No. 1.3e-33;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFP 60
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFK 60

Qy 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 99
Db 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 99
```

```
US-08-529-055-50

Query Match      79.4%; Score 390; DB 4; Length 183;
Best Local Similarity 81.0%; Pred. No. 7.7e-34;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFP 60
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFQ 60

Qy 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 100
Db 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 100

RESULT 10
US-08-529-055-64
; Sequence 64, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-529-055-64

Query Match      79.0%; Score 388; DB 4; Length 194;
Best Local Similarity 81.8%; Pred. No. 1.3e-33;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFP 60
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDKLDLDAEIAKLEKVDGDFK 60

Qy 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 99
Db 61 NSDGEQAGQYLVAEKDLDKAEALGNTGADLKKAVIDEPE 99
```



## RESULT 11

US-08-710-749-4  
; Sequence 4, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 4:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-4

Query Match 78.7%; Score 386.5; DB 2; Length 101;  
Best Local Similarity 82.2%; Pred. No. 8.7e-34;  
Matches 83; Conservative 6; Mismatches 11; Indels 1; Gaps 1;  
QY 1 LKEIDESSEDYEKEGLRAPLQSKL-DAKKAKLSKLDXSDKXDELDAEIAKLEKDVGF 59  
Db 1 LKEIDESSEDYVKEGERAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 60  
QY 60 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100  
Db 61 KNSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAVIDEPE 101

## RESULT 12

US-09-583-110-4871  
; Sequence 4871, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4871  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4871

Query Match 77.8%; Score 382; DB 4; Length 550;  
Best Local Similarity 80.0%; Pred. No. 2.1e-32;  
Matches 80; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
QY 1 LKEIDESSEDYEKEGLRAPLQSKLDAKAKLSKLDXSDKXDELDAEIAKLEKDVGF 60  
Db 144 LKEIDESSEDYVKEGLRAPLQSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 203  
QY 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVIDEPE 100  
Db 204 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVIDEPE 243

## RESULT 13

US-09-107-433-3858  
; Sequence 3858, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
; THERAPEUTICS  
; NUMBER OF SEQUENCES: 5206  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: <Unknown>  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: <Unknown>  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,433  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/ 085131  
; FILING DATE: May 12, 1998  
; APPLICATION NUMBER: 60/051553  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneka  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 3858:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 550 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...550  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:  
US-09-107-433-3858

Query Match 77.8%; Score 382; DB 4; Length 550;  
Best Local Similarity 80.0%; Pred. No. 2.1e-32;  
Matches 80; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDSESDSEYKEGLRAPLQSKLDAKAKLSKLDEXSDKXDELDAEIAKLEKDVGFDP 60  
|||||  
Db 144 LKEIDSESDSEYKEGLRAPLQSKLDAKAKLSKLDEXSDKXDELDAEIAKLEKDVGFDP 60  
|||||  
Qy 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100  
|||||  
Db 204 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243  
|||||

## RESULT 14

US-08-710-749-5

; Sequence 5, Application US/08710749

; Patent No. 5955089

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: Hollingshead, Susan

; APPLICANT: Becker, Robert

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris &amp; Safford

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/710,749

; FILING DATE: 20-SEP-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2074

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 101 amino acids

; TYPE: amino acid

; STRANDEDNESS: n/a

; TOPOLOGY: linear

; MOLECULE TYPE: amino acid

; US-08-710-749-5

Query Match 77.5%; Score 380.5; DB 2; Length 101;  
Best Local Similarity 83.2%; Pred. No. 3.7e-33;  
Matches 84; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDSESDSEYKEGLRAPLQSKL-DAKAKLSKLDEXSDKXDELDAEIAKLEKDVGFDP 59  
|||||  
Db 1 LKEIDSESDSEYKEGLRAPLQSKLSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 60  
|||||  
Qy 60 PMSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100  
|||||  
Db 61 KNSDGEQAGQYLAAAEEDLIAKXAEADLKKAVDEPE 101  
|||||

## RESULT 15

US-09-147-875A-2

; Sequence 2, Application US/09147875A

; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; US-09-147-875A-2

Query Match 77.4%; Score 380; DB 4; Length 100;  
Best Local Similarity 80.0%; Pred. No. 4.2e-33;  
Matches 80; Conservative 7; Mismatches 13; Indels 0; Gaps 0;  
Qy 1 LKEIDSESDSEYKEGLRAPLQSKLDAKAKLSKLDEXSDKXDELDAEIAKLEKDVGFDP 60  
|||||  
Db 1 LKEIDSESDSEYKEGLRAPLQSKLSELDAAKQAKLSKLEELSDKIDELDAEIAKLEKDVDF 60  
|||||  
Qy 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100  
|||||  
Db 61 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVHEPE 100  
|||||

Search completed: June 18, 2005, 17:07:07

Job time : 19.9189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-7

Perfect score: 491

Sequence: 1 LKEIDESSEDEYEKGLRAP.....KEAELGNTGADLKKAVDPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	99.2	100	15	US-10-674-755-7
2	479	97.6	166	15	US-10-299-636-63
3	395	80.4	100	15	US-10-674-755-5
4	394	80.2	168	15	US-10-299-636-70
5	391	79.6	100	15	US-10-674-755-6
6	390	79.4	100	15	US-10-674-755-4
7	390	79.4	183	15	US-10-299-636-65
8	388	79.0	194	15	US-10-299-636-79
9	380	77.4	100	15	US-10-674-755-2
10	378	77.0	98	15	US-10-674-755-1
11	370	75.4	100	15	US-10-674-755-3
					Sequence 7, Appli
					Sequence 63, Appl
					Sequence 5, Appli
					Sequence 70, Appl
					Sequence 6, Appli
					Sequence 65, Appl
					Sequence 79, Appl
					Sequence 2, Appli
					Sequence 1, Appli
					Sequence 3, Appli

12	341.5	69.6	185	15	US-10-299-636-61	Sequence 61, Appl
13	339.5	69.1	101	15	US-10-674-755-9	Sequence 9, Appli
14	336	68.4	100	15	US-10-674-755-8	Sequence 8, Appli
15	297.5	60.6	170	15	US-10-299-636-75	Sequence 75, Appl
16	297.5	60.6	181	15	US-10-299-636-57	Sequence 57, Appl
17	297.5	60.6	643	15	US-10-299-636-95	Sequence 95, Appl
18	297.5	60.6	670	9	US-09-748-875-63	Sequence 63, Appl
19	297.5	60.6	670	10	US-09-298-523B-63	Sequence 63, Appl
20	297.5	60.6	690	9	US-09-748-875-61	Sequence 61, Appl
21	297.5	60.6	690	10	US-09-298-523B-61	Sequence 61, Appl
22	297.5	60.6	691	9	US-09-748-875-1	Sequence 1, Appli
23	297.5	60.6	691	10	US-09-298-523B-1	Sequence 1, Appli
24	297.5	60.6	701	9	US-09-748-875-62	Sequence 62, Appl
25	297.5	60.6	701	10	US-09-298-523B-62	Sequence 62, Appl
26	297.5	60.6	707	9	US-09-748-875-2	Sequence 2, Appli
27	297.5	60.6	707	10	US-09-298-523B-2	Sequence 2, Appli
28	297.5	60.6	711	9	US-09-748-875-3	Sequence 3, Appli
29	297.5	60.6	711	10	US-09-298-523B-3	Sequence 3, Appli
30	297.5	60.6	739	17	US-10-732-923-3294	Sequence 3294, Ap
31	297.5	60.6	929	9	US-09-748-875-60	Sequence 60, Appl
32	297.5	60.6	929	10	US-09-298-523B-60	Sequence 60, Appl
33	297.5	60.6	929	15	US-10-299-636-94	Sequence 94, Appl
34	295	60.1	100	15	US-10-674-755-12	Sequence 12, Appl
35	294.5	60.0	188	15	US-10-299-636-74	Sequence 74, Appl
36	291.5	59.4	198	15	US-10-299-636-105	Sequence 105, App
37	291.5	59.4	354	15	US-10-299-636-96	Sequence 96, Appl
38	291.5	59.4	588	15	US-10-299-636-96	Sequence 96, Appl
39	291.5	59.4	619	10	US-09-882-774-1	Sequence 1, Appli
40	291.5	59.4	619	15	US-10-282-122A-73702	Sequence 73702, A
41	291.5	59.4	619	16	US-10-414-532-72	Sequence 72, Appl
42	290.5	59.2	99	15	US-10-674-755-16	Sequence 16, Appl
43	289	58.9	100	15	US-10-674-755-11	Sequence 11, Appl
44	286.5	58.4	99	15	US-10-674-755-11	Sequence 11, Appl
45	286.5	58.4	204	15	US-10-299-636-66	Sequence 66, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-674-755-7  
; Sequence 7, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-7

Query Match 99.2%; Score 487; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2e-37;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDESSEDEYEKGLRAPLQSKLDKAKLKLKLDKSDKXDKDLDAETAKLEKDVGDPP 60  
DB 1 LKEIDESSEDEYEKGLRAPLQSKLDKAKLKLKLDKSDKXDKDLDAETAKLEKDVGDPP 60  
QY 61 NSDSEQAQGYLVAAEKLDLDAKEAELGNTGADLKKAVDPE 100  
|||||

Db 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100

## RESULT 2

US-10-299-636-63  
; Sequence 63, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 63  
; LENGTH: 166  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-63

Query Match 97.6%; Score 479; DB 15; Length 166;  
Best Local Similarity 97.0%; Pred. No. 1.9e-36;  
Matches 97; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Db 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Qy 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100

Db 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100

## RESULT 3

US-10-674-755-5  
; Sequence 5, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-5

Query Match 80.4%; Score 395; DB 15; Length 100;  
Best Local Similarity 83.0%; Pred. No. 6e-29; Mismatches 5; Indels 0; Gaps 0;  
Matches 83; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Db 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Qy 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100  
Db 61 NSDGEQAGQYLVAAEEDLIAKKAELEQTEADLKKAVHEPE 100

## RESULT 4

US-10-299-636-70  
; Sequence 70, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa at position 38 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (42)  
; OTHER INFORMATION: Xaa at position 42 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (84)  
; OTHER INFORMATION: Xaa at position 84 is unknown  
; OTHER INFORMATION: Xaa at position 84 is unknown  
US-10-299-636-70

Query Match 80.2%; Score 394; DB 15; Length 168;  
Best Local Similarity 84.0%; Pred. No. 1.3e-28;  
Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Db 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKSDKXDLDAEIAKLEKXGVDFP 60

Qy 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100

Db 61 NSDGEQAGQYLVAAEEDLIAKKAELEQTEADLKKAVHEPE 100

## RESULT 5

US-10-674-755-6  
; Sequence 6, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
US-10-674-755-6

```
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6

Query Match          79.6%; Score 391; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 1.4e-28;
Matches 84; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKVDGDP 60
    |||||
Db 1 LKEIDSDSEYKGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVDGDP 60
    |||||
QY 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100
    |||||
Db 61 NSDGEQAGQYLAAAEEDLIAXKAELEKAEADLKKAVDEPE 100
    |||||

RESULT 6
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match          79.4%; Score 390; DB 15; Length 100;
Best Local Similarity 82.0%; Pred. No. 1.7e-28;
Matches 82; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKVDGDP 60
    |||||
Db 1 LKEIDSDSEYKGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVDGDP 60
    |||||
QY 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 100
    |||||
Db 61 NSDGEQAGQYLAAAEEDLIAXKAELEKAEADLKKAVDEPE 100
    |||||

RESULT 7
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match          79.0%; Score 388; DB 15; Length 194;
Best Local Similarity 81.8%; Pred. No. 5.7e-28;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSDSEYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKVDGDP 60
    |||||
Db 1 LKEIDSDSEYKGERAPLQSELDKQAKLSKLEESDKXDELDAETAKLEKVDGDP 60
    |||||
QY 61 NSDGEQAGQYLVAAEKDLDAKEAELGNTGADLKKAVDEPE 99
    |||||
Db 61 XSDGEQAGQYLAAAEEDLIAXKAELEQTEADLKKAVNEP 99
    |||||

RESULT 9
US-10-674-755-2
```

```
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match      77.4%; Score 380; DB 15; Length 100;
Best Local Similarity 80.0%; Pred. No. 1.4e-27;
Matches 80; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 NSDGEAQYLVAAEKDLDAKAEELGNTGADLKKAHVDEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSNGEAEQYRAAEEDLAAKQAELEKTEADLKKAHVHEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match      77.0%; Score 378; DB 15; Length 98;
Best Local Similarity 83.0%; Pred. No. 2.2e-27;
Matches 83; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 NSDGEAQYLVAAEKDLDAKAEELGNTGADLKKAHVDEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSDGEQA-QYLAANEEDL-AKAELEKTEADLKKAHVHEPE 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-674-755-3
; Sequence 3, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      75.4%; Score 370; DB 15; Length 100;
Best Local Similarity 78.0%; Pred. No. 1.2e-26;
Matches 78; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 NSDGEAQYLVAAEKDLDAKAEELGNTGADLKKAHVDEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NSNGEAEQYRAAEEDLAAKQAELEKTEADLKKAHVHEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-61

Query Match      69.6%; Score 341.5; DB 15; Length 185;
Best Local Similarity 74.3%; Pred. No. 1e-23;
Matches 75; Conservative 7; Mismatches 18; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 59
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LKEIDESDSEYKEGLRAPLQSKLDKAKKLSKLDKXSKKLDLDAEIAKLEKDVGFDP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 60 PMSDGEAQYLVAAEKDLDAKAEELGNTGADLKKAHVDEPE 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 QNSGGYGYLYLEAREKDLVAKKAELEKTEADLKKAHVHEPE 101
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-674-755-9
; Sequence 9, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
```

```

; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match          60.6%; Score 297.5; DB 15; Length 170;
Best Local Similarity 67.0%; Pred.No.1.1e-19;
Matches 67; Conservative 7; Mismatches 25; Indels 1; Gaps 1

Qy 1 LKEIDESDSDEYKEGLRAPLQSKLDAAKAKLSKLDKSKDKKDELDAEIAKLEKDVGDFF 60
    |||||
Db 1 LKEIDESDSDEYKGLRAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60
    |||||

Qy 61 NSDGEQAGQYLVAAEKDLDAAKEAELGNTGADLKKAVDEPE 100
    :: : : |||
Db 61 GNNNVEA-YFKEGLEKTTAEKKAEELEKAEADLKKAVDEPE 99
    |||

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-7  
Perfect score: 491  
Sequence: 1 LKEIDSESDYEKEGLRAP.....KEAELGNTGADLKKAVDPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	291.5	59.4	619	2 A97887	surface protein ps
2	291.5	59.4	619	2 A41971	surface protein ps
3	100	20.4	744	2 F95013	pneumococcal surfa
4	93	18.9	405	2 A3939	Fc gamma (IgG) rec
5	92	18.7	1963	1 MWRW	myosin heavy chain
6	91	18.5	387	2 S57834	fcra protein precu
7	90	18.3	896	2 S43074	epidermal growth f
8	89	18.1	166	2 S73342	hypothetical prote
9	89	18.1	388	2 A46173	hypothetical prote
10	88	17.9	1006	2 C70445	Mip4 protein - Str
11	88	17.9	1138	2 T24635	ATPase subunit of
12	88	17.9	3488	2 T34418	hypothetical prote
13	86.5	17.6	501	2 A44643	hypothetical prote
14	86	17.5	532	2 S54871	M protein precuso
15	86	17.5	886	2 H63378	conserved hypotnet
16	85.5	17.4	577	2 S30237	transcription init
17	85.5	17.4	668	2 B86831	hypothetical prote
18	85	17.3	292	2 I52858	nucleolar protein
19	84.5	17.2	369	2 AG1648	hypothetical prote
20	84.5	17.2	837	2 D69171	intracellular prot
21	84.5	17.2	1177	2 B75150	chromosome segrega
22	84	17.1	161	2 S48396	tropomyosin TPM2 -
23	84	17.1	454	2 S43556	plasmaingogen-bindin
24	84	17.1	522	2 C96608	hypothetical prote
25	84	17.1	1938	1 A40997	myosin heavy chain
26	83.5	17.0	1034	2 T32297	hypothetical prote
27	83	16.9	213	2 S54579	hypothetical prote
28	83	16.9	897	2 A54696	EGF receptor subst
29	82.5	16.8	764	2 T05409	hypothetical prote

30	82.5	16.8	1421	2 T05892	hypothetical prote
31	82	16.7	257	2 A34168	nucleolar phosphop
32	82	16.7	281	2 F75216	hypothetical prote
33	82	16.7	292	2 A36089	nucleolar phosphop
34	82	16.7	292	2 A28939	nucleophosmin - ra
35	82	16.7	388	2 S52536	fcra 15 protein -
36	82	16.7	518	2 G84488	En/Spm-like transp
37	82	16.7	539	2 A28549	M24 protein precu
38	82	16.7	557	2 A96527	probable nuM1 prot
39	82	16.7	746	2 T47237	myosin II heavy ch
40	82	16.7	752	2 G96510	hypothetical prote
41	81.5	16.6	880	2 F75103	conserved hypotnet
42	81	16.5	231	2 S0460	hypothetical prote
43	81	16.5	357	2 JC4090	FK506-binding 39k
44	81	16.5	472	2 S43554	plasmaingogen-bindin
45	81	16.5	780	2 F96840	hypothetical prote

ALIGNMENTS

RESULT 1

A97887  
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.,  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID:91  
C:Genetics:  
A:Gene: pspA

Query Match	59.4%	Score	291.5;	DB	2;	Length	619;
Best Local Similarity	65.0%	Pred. No.	1.2e-15;				
Matches	65;	Conservative	10;	Mismatches	24;	Indels	1;
Gaps	1;						
Qy	1	LKEIDSESDYEKEGLRAPLQSKLDAKAKLKLDESDKXDELDAIAKLEKDVGDPE	60				
Db	223	LKEIDSESDYEKEGFRAPLQSKLDAKAKLKLDESDKXDELDAIAKLEKDVGDPE	282				
Qy	61	NSDGEQAGYLVAABKLDAAEALGNTGADLKKAVDPE	100				
Db	283	ENNVE-DYFKEGLEKTTAAKAELEKTEADLKKAVDPE	321				

RESULT 2

A41971  
surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A>Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:g153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

RESULT 4

A33939

Fc gamma (Igg) receptor II precursor - Streptococcus sp. (fragment)

C:Species: Streptococcus sp.

C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 26-Aug-1999

C:Accession: A33939

R:Heath, D.G.; Cleary, P.P.

Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989

A:Title: Fc-receptor and M-protein genes of group A streptococci are products of gene dup-

A:Reference number: A33939; MUID:89282846; PMID:2660147

A:Accession: A33939

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <HEA>

A:Cross-references: GB:M22532; NID:g153628; PIDN:AAB95296.1; PID:g552003

C:Superfamily: M5 protein

C:Keywords: immunoglobulin receptor

Query Match 18.9%; Score 93; DB 2; Length 405;

Best Local Similarity 27.1%; Pred. No. 2.8;

Matches 32; Conservative 23; Mismatches 45; Indels 18; Gaps 3;

Qy 1 LKEIDESPDEYK-EGLRAPLOSXDAKAKL-----SKLDEKSDKXDLDAEIAKLEKD 55

Db 208 LKQDAKTEHIAKLSAATLENLGSAKRELTDLQAKLDTATAEKAKLESQVTTLENL 267

Qy 56 V-----DFPNSDGEQAGQYLVAAEKDLDKAKEALGNLTGADLKXAVDPPE 100

Db 268 LGSAKRELTDLQAKLDAANAEEKQLSQSQAALKEQLATKKELADLQAKLAATNQEKE 325

RESULT 5

MMKW

myosin heavy chain B [similarity] - Caenorhabditis elegans

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Caenorhabditis elegans

C:Date: 13-Jun-1983 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C:Accession: T20770; T21629; A93958; A93287; A21074; A02992

R:Kershaw, J.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19322

A:Accession: T20770

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1963 <WLT>

A:Cross-references: UNIPROT:O02244; EMBL:Z81499; PIDN:CAB05505.1; GSPDB:GN00019; CESP:F11C3.3

A:Experimental source: clone F11C3

A:Accession: T21629

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1963 <WLT2>

A:Cross-references: EMBL:Z83107; PIDN:CAB05505.1; GSPDB:GN00019; CESP:F11C3.3

A:Experimental source: clone F32A7

R:Karn, J.; Brenner, S.; Barnett, L.

Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983

A:Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy cha

A:Reference number: A93958; MUID:83273600; PMID:6576334

A:Accession: A93958

A:Molecule type: DNA

A:Residues: 1-61,'EMSVIQ',65-376,'V',378-1963 <KAR>

A:Cross-references: GB:J01050; NID:g156399; PIDN:AAA28124.1; PID:g156400

R:McLuachlan, A.D.; Karn, J.

Nature 299, 228-231, 1982

A:Title: Periodic charge distributions in the myosin rod amino acid sequence match cross-

A:Reference number: A93287; MUID:82272395; PMID:7202124

A:Accession: A93287

A:Molecule type: DNA

A:Residues: 847-1333,'R',1335-1876,'L',1878-1963 <MCI>

R:Willis, N.; Gesteland, R.F.; Karn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.

Cell 33, 575-583, 1983

A:Title: The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsense

C;Accession: S43074; I38525  
C;Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.  
Oncogene 9, 1039-1045, 1994  
A;Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4,  
A;Reference number: S43074; MUID:94181254; PMID:8134107  
A;Accession: S43074  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-896 <BER>  
A;Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA82305.1; PID:g4701  
R;Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner,  
Oncogene 9, 1591-1597, 1994  
A;Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in ev  
A;Reference number: I38525; MUID:94239734; PMID:8183552  
A;Accession: I38525  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-821,'M',823-896 <RES>  
A;Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260  
C;Genetics:  
A;Gene: GDB:EPS15; AF-1p; MLT5  
A;Cross-references: GDB:360337; OMIM:600051  
A;Map position: lp32-1p32

```

Query Match      18.3%; Score 90; DB 2; Length 896;
Best Local Similarity 22.8%; Pred. No. 11;
Matches 28; Conservative 28; Mismatches 41; Indels 26; Gaps 3;

Qy      1 LKEIDESDE---DYEGLRAPLQSKLDKAKKSLKLDXSKD----- 41
Db      332 IKELDTLNNEIVDLQREKNVQDLKEKEDTTIKRTSEVODLQDEVQRENTNLQKLOAQK 391

Qy      42 -----XBELDAEIAKLEKDVGDFFPNSDGEQAGQYLVAAEKIDIAEALGNTGADLKKA 95
Db      392 QQVOELDELDEQQALESQLEKVRKCAEEA-QLISSLKAELTSQESISTYEELAKA 450

Qy      96 VDE 98
Db      451 REE 453

```

RESULT 8  
S73342  
hypothetical protein E07\_orf166 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S73342  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73342  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-166 #HMM>  
A:Cross-references: UNIPROT:P75260; EMBL:AE000003; GB:U00089; NID:g1673661; PIDN:AAB9566  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3

```

Query Match      18.1%; Score 89; DB 2; Length 166;
Best Local Similarity 36.8%; Pred.No.2.4;
Matches 25; Conservative 13; Mismatches 24; Indels 6; Gaps 2;

QY 21 LOSKIDAKKAKL-SKLDEKSDKDXDELDAAETAKLEKDVGFPP---NSDGEQAGQYLVAAEKD 77
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 95 LSEKQVKLEAKVDKLEE---KVDKLEAKVDKLEEKVDKLEAKVDKGFEMAKILAAINKR 151

```

Qy	78	LDAKEAEL	85
		:	
D <sub>b</sub>	152	LDSIEGRL	159

```
RESULT 9
A46173
Mip4 protein - Streptococcus sp. (group A)
C:Species: Streptococcus sp.
C>Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C:Accession: A46173
R:O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.
Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992
A:Title: Two major classes in the M protein family in group A streptococci.
A:Reference number: A46173; MUID:92409576; PMID:1528877
A:Contents: group A
A:Accession: A46173
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-388 <OIT>
A>Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBIPI:114064)
C:Superfamily: M5 protein

Query Match 18.1%; Score 89; DB 2; Length 388;
Best Local Similarity 26.3%; Pred. No. 5.6;
Matches 31; Conservative 23; Mismatches 46; Indels 18; Gaps 3;

Qy 1 LKEDIQSDSEYDK-EGLRAPLQSKLDKAKL-----SKLDEXDKXDELDAEIAKLEKD 55
Db 171 LKQDASKTEBIAKQSEATLENLGSAKRELTELQKLDATAEAKAKLESQVTTLENL 230

Qy 56 VG-----DFPNSDGEQAGGYLVAAEKDLDAKEAELGNTGADLKKAVIDEPE 100
Db 231 LGSAKRELTDLQKLDAAAEKELKQSQATLEKLEATKELADLQAKLAATNQEK 288

RESULT 10
C70445
ATPase subunit of ATP-dependent proteinase (BC 3.4.-.-) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70445
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70445
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1006 <AOQ>
A:Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298
A:Experimental source: strain VP5
C:Genetics:
A:Gene: clpB
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: hydrolase

Query Match 17.9%; Score 88; DB 2; Length 1006;
Best Local Similarity 33.0%; Pred. No. 17;
Matches 35; Conservative 16; Mismatches 35; Indels 20; Gaps 4;

Qy 1 LKEDIQSDSE-----DYEKEGLRAPLQSKLDKAKLKS-LDEXSDKXDELDAEIAKLEK 54
Db 552 IKALEEQIIEANLKGDIYE-----AQLKIEKAKLEKEQELLGKVGVEAKLEKK 604

Qy 55 DVGDFPNSDGEQAGGYLVAAEKDLDAKEAELGNTGADLKKAVIDEPE 100
Db 605 KIEELDEKIKE-----AAEKGYDYKEAELKIEKAKLEKELKLE 643

RESULT 11
T24635
hypothetical protein T07C4.10a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
```

```
C:Accession: T24635; S41023
R:Buck, D.
submitted to the EMBL Data Library, February 1995
A:Reference number: Z19915
A:Accession: T24635
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1138 <WIL>
A:Cross-references: UNIPROT:Q22276; EMBL:Z48055; PIDN:CAA88136.1; GSPDB:GN00021; CESP:TO:
A:Experimental source: clone T07A5
R:Barks, M.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41014
A:Accession: S41023
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 895-1138 <BER>
A:Cross-references: EMBL:Z29443
C:Genetics:
A:Gene: CESP:T07C4.10a
A:Map position: 3
A:Introns: 16/1; 124/3; 146/3; 204/1; 311/3; 358/3; 615/3; 900/2; 961/3; 1008/2; 1110/3

Query Match 17.9%; Score 88; DB 2; Length 1138;
Best Local Similarity 24.8%; Pred. No. 20;
Matches 33; Conservative 21; Mismatches 41; Indels 38; Gaps 4;

Qy 3 EIDESDSE-----DYK-----EGLRAPLQSKLDKAKL----- 32
Db 359 EVDQLHSEIEVIGKKKSDLENRLFYDEKLRQAQFQDENKLRADLEKKLKTSQEKLVKYE 418

Qy 33 -----SKLDEXDKXDELDAEIAKL--EKDVGDFPNSDGEQAGGYLVAAEKDLDAKEA 84
Db 419 KIBELQSLRLNKKELEEVQAEKNKLLDKNTDFELDEAKVQGEHLEKORKEAWEKVEQ 478

Qy 85 LGNTGADLKKAVID 97
Db 479 LQEMLGELEELD 491

RESULT 12
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34418
R:Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: Z21521
A:Accession: T34418
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3488 <FUL>
A:Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
A:Experimental source: strain Bristol N2; clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.3
A:Map position: 5
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 17.9%; Score 88; DB 2; Length 3488;
Best Local Similarity 29.5%; Pred. No. 60;
Matches 31; Conservative 23; Mismatches 43; Indels 8; Gaps 5;

Qy 1 LKEDIQSDSEYDK-EGLRAPLQSKLDKAKLKS--KLDEXSDKXDELDAEIAKLEKDV 57
Db 1041 IKKVEDDAARKEKELNDKLESEIATKKSADKLEEQAKAAKAAVEAAKKQEKD 1100

Qy 58 DFPNSDGEQAGGYLVAAEKDLDAKEAELGN-TGADL---KKAVDE 98
Db 1101 EQLKLDTEAASKKAAAEKLELE-KQAQIKKAAGADAVKKQKELDE 1144
```

## RESULT 13

A44643  
M protein precursor - Streptococcus pyogenes (serotype M57) (fragment)  
C:Species: Streptococcus pyogenes  
A:Variety: serotype M57  
C:Date: 26-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
C:Accession: A44643; S60833  
R:Manjula, B.N.; Khandke, K.M.; Fairwell, T.; Relf, W.A.; Sriprakash, K.S.  
J. Protein Chem. 10, 369-384, 1991  
A:Title: Heptad motifs within the distal subdomain of the coiled-coil rod region of M protein  
A:Note: nucleotide sequence of the M57 gene and relation of the deduced  
A:Reference number: A44643; PMID:92143933; PMID:1781883  
A:Accession: A44643  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-501 <MAN>  
A:Experimental source: type M57, strain A995  
A:Note: sequence inconsistent with nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:83737, NCBI:83738)  
A:Note: parts of this sequence were confirmed by peptide sequencing  
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop  
A:Reference number: S60784; PMID:95198537; PMID:7891551  
A:Accession: S60833  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 14-95 <WHA>  
A:Cross-references: EMBL:U11971  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Superfamily: M5 protein  
C:Keywords: coiled coil; dimer

Query Match 17.6%; Score 86.5; DB 2; Length 501;  
Best Local Similarity 27.4%; Pred. No. 11;  
Matches 34; Conservative 20; Mismatches 37; Indels 33; Gaps 4;  
QY 2 KEIDSESDYEKEGLRAPLQSKLDAAKALSK-----LDXSDKXD 43  
DB 234 RELTELQAKLDETQKELANQAQLDAYKAEIAKLQENKISEASRQGLRRDLDSREAKK 293  
QY 44 ELDAETAKLEKDVGFPPNSDGGQAGQYLVAASEKDLDAK-----EALNGTGADLKAV 96  
DB 294 QLEAEHQKLEEQ-----NKISEASRQGL---RRDLDSREAKKQVEKDLANLTAEIDKVK 345  
QY 97 DEPE 100  
DB 346 BEKQ 349

## RESULT 14

S54871  
M protein - Streptococcus sp.  
C:Species: Streptococcus sp.  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: S54871  
R:Podbielaki, A.; Melzer, B.  
submitted to the EMBL Data Library, June 1991  
A:Reference number: S54871  
A:Accession: S54871  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-532 <POD>  
A:Cross-references: UNIPROT:Q55098; EMBL:X60097; NID:9840905; PIDN:CAA42693.1; PID:98409  
C:Superfamily: M5 protein

Query Match 17.5%; Score 86; DB 2; Length 532;  
Best Local Similarity 26.6%; Pred. No. 13;  
Matches 33; Conservative 17; Mismatches 38; Indels 36; Gaps 5;  
QY 5 DSDSESDYEKEGLRAPLQSKLDAAK---AKSKLDEKSDKXD----- 43

DB 253 BEKQISDASRQSLRRDLDSREAKKQLEAEYQKLEEEKQISDASRQSLRRDLDSREAKK 312  
QY 44 ELDAETAKLEKDVGFPPNSDGGQAGQYLVAASEKDLDAK-----EALNGTGADLKAV 96  
DB 313 QLEAEHQKLEEQ-----NKISEASRQGL---RRDLDSREAKKQVEKDLANLTAEIDKVK 364  
QY 97 DEPE 100  
DB 365 BEKQ 368

## RESULT 15

H69378  
conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: H69378  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A:Reference number: A69250; PMID:98049343; PMID:9389475  
A:Accession: H69378  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-886 <KLR>  
A:Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:92689355; PIDN:AAB9031  
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 17.5%; Score 86; DB 2; Length 886;  
Best Local Similarity 23.9%; Pred. No. 22;  
Matches 32; Conservative 26; Mismatches 40; Indels 36; Gaps 3;  
QY 1 LKEDSESDYEKEG-----LRAPLQSKLDAAKAKLKLDEKSDKXDLDAEIAKLEK-- 54  
DB 296 LSEINQALRDVKEKRGDUTREAAAGIQALKKAEDNSKLEETTKRIEELERLEPERFKSH 355  
QY 55 -----DVGDFPN-----SDGEQAGQYLVAASEKDLDAKAE 84  
DB 356 RLLETLPKPMRMOGIKAKLEEKNLTPDKVEKMYDLSKAKEBEKEITEKLLKLIACKSS 415  
QY 85 LGNTGADLKKAVIDE 98  
DB 416 LKTRGAQLKKAVIDE 429

Search completed: June 18, 2005, 17:03:55  
Job time : 15.113 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)

840.012 Million cell updates/sec

Title: US-10-674-755-7

Perfect score: 491

Sequence: 1 LKEIDSESDYEKEGLRAP.....KEALNGTGADLKKAVIDEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	427	87.0	194	2 Q9L5B5	Q9L5B5 streptococc
2	427	87.0	218	2 Q6UEB2	Q6UEB2 streptococc
3	427	87.0	233	2 Q9L568	Q9L568 streptococc
4	427	87.0	236	2 Q9L569	Q9L569 streptococc
5	427	87.0	243	2 Q9L564	Q9L564 streptococc
6	427	87.0	243	2 Q9L567	Q9L567 streptococc
7	427	87.0	244	2 Q9L565	Q9L565 streptococc
8	427	87.0	247	2 Q9L566	Q9L566 streptococc
9	427	87.0	249	2 Q9L570	Q9L570 streptococc
10	427	87.0	254	2 Q9L563	Q9L563 streptococc
11	427	87.0	401	2 Q9LAZ2	Q9LAZ2 streptococc
12	400	81.5	394	2 Q9LAY6	Q9LAY6 streptococc
13	400	81.5	395	2 Q9LAZ1	Q9LAZ1 streptococc
14	400	81.5	406	2 Q9LAZ0	Q9LAZ0 streptococc
15	394	80.2	340	2 Q8KQK5	Q8KQK5 streptococc
16	383	78.0	225	2 Q9L591	Q9L591 streptococc
17	382	77.8	222	2 Q9L577	Q9L577 streptococc
18	382	77.8	262	2 Q9L576	Q9L576 streptococc
19	382	77.8	415	2 Q9LAY7	Q9LAY7 streptococc
20	378	77.0	246	2 Q9L578	Q9L578 streptococc
21	376	76.6	255	2 Q9L581	Q9L581 streptococc
22	376	76.6	255	2 Q9L5B6	Q9L5B6 streptococc
23	375	76.4	416	2 Q9LAY8	Q9LAY8 streptococc
24	356	72.5	393	2 Q9LAZ3	Q9LAZ3 streptococc
25	347	70.7	207	2 Q8GNS9	Q8GNS9 streptococc
26	346	70.5	237	2 Q9L592	Q9L592 streptococc
27	346	70.5	395	2 Q9LAY9	Q9LAY9 streptococc
28	297.5	60.6	739	2 Q9RQT4	Q9RQT4 streptococc
29	297.5	60.6	820	2 Q9RQT1	Q9RQT1 streptococc
30	297.5	60.6	929	2 Q9KK19	Q9KK19 streptococc
31	297.5	60.6	929	2 Q9ZAY5	Q9ZAY5 streptococc

32 291.5 59.4 619 2 Q54972 Q54972 streptococc  
33 291.5 59.4 619 2 Q8DRI0 Q8DRI0 streptococc  
34 288.5 58.8 415 2 Q9LAY1 Q9LAY1 streptococc  
35 286.5 58.4 417 2 Q9LAY3 Q9LAY3 streptococc  
36 285.5 58.1 437 2 Q9LAY4 Q9LAY4 streptococc  
37 269.5 54.9 395 2 Q9LAY2 Q9LAY2 streptococc  
38 269.5 54.9 408 2 Q9LAY0 Q9LAY0 streptococc  
39 268.5 54.7 99 2 Q8KQK4 Q8KQK4 streptococc  
40 268.5 54.7 249 2 Q9L575 Q9L575 streptococc  
41 265.5 54.1 426 2 Q9LAY5 Q9LAY5 streptococc  
42 262 53.4 869 2 Q9KK27 Q9KK27 streptococc  
43 261.5 53.3 224 2 Q8GNS8 Q8GNS8 streptococc  
44 147.5 30.0 479 2 Q9LAX2 Q9LAX2 streptococc  
45 147.5 30.0 480 2 Q9LAX3 Q9LAX3 streptococc

#### ALIGNMENTS

RESULT 1  
Q9L5B5 PRELIMINARY; PRT; 194 AA.  
AC Q9L5B5;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP196;  
RX MEDLINE=20472598; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones";  
RL J. Clin. Microbiol. 38:3663-3669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP196;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AF253407; RAP67355.1;  
DR InterPro; IPR009082; His\_kin\_homodim.  
FT NON\_TER 1 194  
SQ SEQUENCE 194 AA; 21116 MW; E68189FCA2B244F8 CRC64;

Query Match 87.0%; Score 427; DB 2; Length 194;

Best Local Similarity 89.0%; Pred. No. 7.3e-24;  
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 LKEIDSESDYEKEGLRAPLQSKLDKAKLSDKXDKDLDAETAKLEKVDGDFP 60  
|||||

Db 55 LKEIDSESDYEKEGLRAPLQSKLDKAKLSELSDKIDELDAETAKLEKVDGDFK 114  
|||||

QY 61 NSDGSQAQYLVAAFKDLDAKAEALNGTGADLKKAVIDEPE 100  
|||||

Db 115 NSDGSQAQYLVAAFKDLDAKAEALNGTGADLKKAVIDEPE 154  
|||||

RESULT 2

Q6UEB2 PRELIMINARY; PRT; 218 AA.  
AC Q6UEB2;  
DT 05-JUL-2004 (Tremblrel. 27, Created)

DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE PspA (Fragment).





```

J. Clin. Microbiol. 38:3663-3669(2000).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70093.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDVGF 60
Db 74 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 133
Qy 61 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 134 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 173

RESULT 6
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255905; AAF70095.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 244
FT NON_TER 244
SQ SEQUENCE 244 AA; 25946 MW; F9274FFD1957DD06 CRC64;

Query Match 87.0%; Score 427; DB 2; Length 244;
Best Local Similarity 89.0%; Pred. No. 9.1e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 60
Db 51 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 150
Qy 61 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 111 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 150

RESULT 7
Q9L565
ID Q9L565 PRELIMINARY; PRT; 244 AA.
AC Q9L565;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255906; AAF70096.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 60
Db 74 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 133
Qy 61 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 134 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 173

RESULT 8
Q9L566
ID Q9L566 PRELIMINARY; PRT; 247 AA.
AC Q9L566;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255905; AAF70095.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 60
Db 50 LKEIDSDSDYKEGLRAPLQSKLDKAKLSDKXDELDAAIAKLEKDV 109
Qy 61 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 110 NSDGEQAGQYLVAEKDLDKAEKELGNTGADLKKAVIDEPE 149

```

```

J. Clin. Microbiol. 38:3663-3669(2000).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70093.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 60
Db 74 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 133
Qy 61 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 134 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 173

RESULT 6
Q9L567
ID Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255905; AAF70095.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 244
FT NON_TER 244
SQ SEQUENCE 244 AA; 25946 MW; F9274FFD1957DD06 CRC64;

Query Match 87.0%; Score 427; DB 2; Length 244;
Best Local Similarity 89.0%; Pred. No. 9.1e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 60
Db 51 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 150
Qy 61 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 111 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 150

RESULT 7
Q9L565
ID Q9L565 PRELIMINARY; PRT; 244 AA.
AC Q9L565;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255906; AAF70096.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 60
Db 74 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 133
Qy 61 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 134 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 173

RESULT 8
Q9L566
ID Q9L566 PRELIMINARY; PRT; 247 AA.
AC Q9L566;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR EMBL; AF255905; AAF70095.1; -
DR InterPro; IPR009082; His_kin_homodim.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
FT NON_TER 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match 87.0%; Score 427; DB 2; Length 243;
Best Local Similarity 89.0%; Pred. No. 9e-24;
Matches 89; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 60
Db 50 LKEIDSDSDYKEGLRAPLQSKLDKAKKLSKLDKXSDKXDELDAETAKLEKDVGFDP 109
Qy 61 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 100
Db 110 NSDGEAQGYLVAAEKDLDKAEKELGNTGADLKKAVIDEPE 149

```





```

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=papA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;
RA DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PapA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082387; AAL92492.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR00533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 340
SQ SEQUENCE 340 AA; 38023 MW; E07ECF00B1FBD57 CRC64;

Query Match      80.2%; Score 394; DB 2; Length 340;
Best Local Similarity 82.0%; Pred. No. 3.2e-21;
Matches 82; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSESDYKVEGLRAPLOSXLDAKAKLSKLDXSDKXDELDAEIAKLEKDVGFPP 60
Db 197 LKEIDSESDYKVEGLRAPLQFELDVKQAKLSKLELSDKIDELDAEIAKLEKDVDFK 256

QY 61 NSDGEQAGQYLVAEKDLDAKEAELGNTGADLKKAVDPE 100
Db 257 NSDGEQAGQYLAAAEEDLVAKKAELKTEADLKKAVDPE 296

```

Search completed: June 18, 2005, 17:01:36  
Job time : 61.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-8

Perfect score: 494

Sequence: 1 LKIDSDSDYVKEGLRAP.....KKALEKTEADLKAVHEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseqp16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	399	80.8	550	8	Adk48356 Streptoco
2	399	80.8	550	8	AdR95223 Novel S.
3	386	78.1	194	2	Aaw14584 Streptoco
4	386	78.1	194	7	Abw02618 Db16ac pn
5	386	78.1	891	6	Abu08487 S. pneumo
6	372	75.3	183	2	Aaw14570 Streptoco
7	372	75.3	183	7	Abw02604 Bg9739c p
8	370	74.9	168	7	Abw02609 L81905c p
9	360.5	73.0	167	2	Aaw14575 Streptoco
10	351	71.1	166	2	Aaw14568 Streptoco
11	351	71.1	166	7	Abw02602 Bg8743c p
12	339.5	68.7	185	2	Aaw14586 Streptoco
13	339.5	68.7	185	7	Abw02600 Ac94c pne
14	320.5	64.9	204	2	Aaw14571 Streptoco
15	320.5	64.9	204	7	Abw02605 Ef1019c p
16	310.5	62.9	170	7	Abw02614 Rct135c p
17	310.5	62.9	181	7	Abw02596 0922134c
18	310.5	62.9	865	6	Abu08489 S. pneumo
19	310.5	62.9	929	2	Aaw14593 Streptoco
20	310.5	62.9	929	2	Aay43384 S. pneumo
21	309.5	62.7	198	7	Abw02615 Rx1c pneu
22	309.5	62.7	315	2	Aay04375 Streptoco
23	309.5	62.7	619	2	Aar63437 Pneumococ
24	309.5	62.7	619	2	Aar87598 Pneumococ
25	309.5	62.7	619	2	Aar86911 Pneumococ

## ALIGNMENTS

### RESULT 1

ADK48356

ID ADK48356 standard; protein; 550 AA.

AC ADK48356;

XX

XX

DT 20-MAY-2004 (first entry)

XX

DE Streptococcus pneumoniae protein, Seq ID No 4871.

XX

KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

XX

OS Streptococcus pneumoniae.

XX

PN US6699703-B1.

XX

PD 02-MAR-2004.

XX

PF 26-MAY-2000; 2000US-00583110.

XX

PR 02-JUL-1997; 97US-0051553P.

XX

PR 12-MAY-1998; 98US-0085131P.

XX

PR 30-JUN-1998; 98US-00107433.

XX

(GENO-) GENOME THERAPEUTICS CORP.

XX

PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;

XX

DR WPI; 2004-212399/20.

XX

DR N-PSDB; ADK45695.

XX

XX

PT New nucleic acid molecules and polypeptides useful for diagnosing,

XX

PT preventing and treating pathological conditions resulting from bacterial

XX

PT infection, e.g. Streptococcus pneumoniae infection, and in drug

XX

PT screening.

XX

XX

XX

XX

XX

XX

XX

XX



```
Query Match          78.1%; Score 386; DB 2; Length 194;
Best Local Similarity 81.8%; Pred. No. 4.2e-28;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKGIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 60
QY 61 LTDAEQTEQYLAAEKDLADKAELEKTEADLKKAVHEP 99
DB 61 XSDGEQAGQYLAAAEEDLIAKKAELEQTEADLKKAVNEP 99

RESULT 4
ABW02618
ID ABW02618 standard; protein; 194 AA.
XX
AC ABW02618;
XX
DT 12-FEB-2004 (first entry)
XX
DE Db16ac pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..194
FT /note= "Xaa = Unknown amino acid"
XX
US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
XX
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
PT Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 64; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 194 AA;

Query Match          78.1%; Score 386; DB 7; Length 194;
Best Local Similarity 81.8%; Pred. No. 4.2e-28;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKGIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 60
DB 1 LKEIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 60
QY 61 LTDAEQTEQYLAAEKDLADKAELEKTEADLKKAVHEP 99
DB 61 XSDGEQAGQYLAAAEEDLIAKKAELEQTEADLKKAVNEP 99

RESULT 5
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
AC ABU08487;
XX
DT 24-JUN-2003 (first entry)
XX
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX antibacterial.
XX
OS Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..8991
FT /note= "All Xaa residues within this sequence are
FT unknown"
XX
US6500613-B1.
XX
PD 31-DEC-2002.
XX
PF 16-SEP-1996; 96US-00714741.
XX
PR 15-SEP-1995; 95US-00529055.
XX
XX (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
XX
XX Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
XX
PT Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
PS Disclosure; Col 145-188; 186pp; English.
XX
CC The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents S. pneumoniae
CC PspA protein
XX
SQ Sequence 8991 AA;

Query Match          78.1%; Score 386; DB 6; Length 8991;
Best Local Similarity 81.8%; Pred. No. 4.3e-26;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKGIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 60
DB 7537 LKEIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVYFK 7596
```

```
QY      61 LTDAEQTEQYLAARAKDLADKAELEKTEADLKKAVHEP 99
DB      7597 XSDGEQAGQYLAARAEEDLIARAELEQTEADLKKAVNEP 7635

RESULT 6
AAW14570
ID      AAW14570 standard; protein; 183 AA.
XX      AC
XX      17-OCT-2003 (revised)
XX      28-OCT-1997 (first entry)
XX      AAW14570;
XX      DE
XX      Streptococcus pneumoniae PspA central region.
XX      KW
XX      PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX      bacteriaemia; pneumonia.
XX      OS
XX      Streptococcus pneumoniae; strain Bg9739.
XX      PN
XX      WO9709994-A1.
XX      PD
XX      20-MAR-1997.
XX      PF
XX      16-SEP-1996; 96WO-US014819.
XX      PR
XX      15-SEP-1995; 95US-00529055.
XX      PA
XX      (UABR-) UAB RES FOUND.
XX      PI
XX      Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
XX      Hollingshead S, Tart R, Brooks-Walter A;
XX      WPI; 1997-202002/18.
XX      DR
XX      Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX      PT
XX      in vaccines for protecting animals against S.pneumoniae infection.
XX      PS
XX      Example 6; Fig 13; 296pp; English.
XX      CC
XX      This sequence shows the central portion, including the C-terminus of the
XX      CC
XX      alpha-helix region and some of the proline-rich region, of pneumococcal
XX      CC
XX      surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.
XX      CC
XX      Comparison of the N-terminal and central regions (AAW14533-57 and
XX      CC
XX      AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX      CC
XX      be used to divide the strains into several families based on sequence
XX      CC
XX      homologies. PspA polypeptides, or fragments of them, can be used in
XX      CC
XX      vaccines to protect animals against S. pneumoniae infection and hence for
XX      CC
XX      the prevention of diseases such as otitis media, meningitis, bacteraemia
XX      CC
XX      and pneumonia. The sequence of the 3' half of the PspA alpha-helical
XX      CC
XX      region and the immediate 5' tip of the coding sequence are likely to be
XX      CC
XX      the critical sequences for predicting PspA cross-reactions and vaccine
XX      CC
XX      composition. (Updated on 17-OCT-2003 to standardise OS field)
XX      CC
XX      Sequence 183 AA;

Query Match      75.3%; Score 372; DB 2; Length 183;
Best Local Similarity 79.0%; Pred. No. 7.9e-27;
Matches 79; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY      1 LKGEIDSESDYVKEGLRAPLQSELDARLTKLSTLELSKIDELDARIPKLEKNVEYFK 60
DB      1 LKGEIDSESDYVKEGLRAPLQSELDARLTKLSTLELSKIDELDARIPKLEKNVEYFK 60

QY      61 LTDAEQTEQYLAARAKDLADKAELEKTEADLKKAVHEP 100
DB      61 NSDGEQAGQYLAARAEEDLIARAELEKAEADLKKAVDEPE 100

RESULT 7
ABW02604
ID      ABW02609 standard; protein; 168 AA.
XX      AC
XX      ABW02609;
XX      12-FEB-2004 (first entry)
XX      DT
XX      XX
```

ABW02604 standard; protein; 183 AA.

ABW02604;

12-FEB-2004 (first entry)

Bg9739c pneumococcal surface protein A (PspA) central region.

Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine; immunological; gene therapy; immunostimulant.

Unidentified.

US6592876-B1.

15-JUL-2003.

15-SEP-1995; 95US-00529055.

20-APR-1993; 93US-00048896.

06-JUN-1995; 95US-00465746.

(UABR-) UAB RES FOUND.

Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A; WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

Example 6; SEQ ID NO 50; 121pp; English.

The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (Pspas) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is Bg9739c pneumococcal surface protein A (PspA) central region. This sequence is used in the exemplification of the invention

Seq Sequence 183 AA;

Query Match 75.3%; Score 372; DB 7; Length 183; Best Local Similarity 79.0%; Pred. No. 7.9e-27; Matches 79; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKGEIDSESDYVKEGLRAPLQSELDARLTKLSTLELSKIDELDARIPKLEKNVEYFK 60  
DB 1 LKGEIDSESDYVKEGLRAPLQSELDARLTKLSTLELSKIDELDARIPKLEKNVEYFK 60  
QY 61 LTDAEQTEQYLAARAKDLADKAELEKTEADLKKAVHEP 100  
DB 61 NSDGEQAGQYLAARAEEDLIARAELEKAEADLKKAVDEPE 100

RESULT 8  
ABW02609  
ID ABW02609 standard; protein; 168 AA.  
XX ABW02609;  
XX 12-FEB-2004 (first entry)  
DT  
XX



L81905c pneumococcal surface protein A (PspA) central region.
Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
immunological; gene therapy; immunostimulant.
Unidentified.
Key Location/Qualifiers
Misc-difference 1..168 /note= "Xaa = Unknown amino acid"
US6592876-B1.
15-JUL-2003.
15-SEP-1995; 95US-00529055.
20-APR-1993; 93US-00048896.
06-JUN-1995; 95US-00465746.
(UABR-) UAB RES FOUND.
Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
WPI; 2003-862841/80.
Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.
Example 6; SEQ ID NO 55; 121pp; English.
The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, an immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is L81905C pneumococcal surface protein A (PepA) central region. This sequence is used in the exemplification of the invention
Seq Sequence 168 AA;
Query Match 74.9%; Score 370; DB 7; Length 168; Best Local Similarity 79.0%; Pred No. 1,le-26; Matches 79; Conservative 4; Mismatches 17; Indels 0; Gaps 0
Gy 1 LKGI DESDSEDYVKEGLRAPLQSELDAKRKTLSLEELSDKDIDLDASPKLENVEYFK 60   :    Db 1 LKEI DESDSEDYVKEGFRAPLQSELDAAQAQLSKLEXSDDKKDLADIASIAKLKVDEDFK 60   :     Gy 61 LTDAEQTEOYLAAAEKOLADKAELEKTEADLLKAVHEPE 100 :  :    Db 61 NSDGEQAGQYLAARERDLIAAKXLKAEARDLKRAVDPE 100   :
RESULT 9
AALW14575 standard; protein; 167 AA. ID AAW14575 XX AC AC AC XX AAW14575; DT 17-OCT-2003 (revised) DT 28-OCT-1997 (first entry)
Streptococcus pneumoniae PepA central region.

```
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Bg8743.
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S., Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8743.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 166 AA;
SQ
Query Match 71.1%; Score 351; DB 2; Length 166;
Best Local Similarity 75.0%; Pred. No. 6.5e-25;
Matches 75; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
Qy 1 LKGI DESDSE DYVKEGLRAPLQSELDAR TKLTLELS DKIDELDAEIPKLEKNVYFK 60
Db 1 LKEI DESDSE DYIKEGLRAPLQSKLDAR KAKLSKLDSELS DKIDELDAEIAKLEKDVGFDP 60
Qy 61 LTDAEQTEQYLAAAEKDLADKKAELEKTEADLKAVHEPE 100
Db 61 NSDGEQAGQYLVAAEKDLDAKEAEELGNTGADLKKAVDEPE 100
RESULT 11
ABW02602
ID ABW02602 standard; protein; 166 AA.
XX
XX ABW02602;
XX
XX 12-FEB-2004 (first entry)
XX
XX Bg8743c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
```

```
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 48; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies), or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bg8743c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX Sequence 166 AA;
SQ
Query Match 71.1%; Score 351; DB 7; Length 166;
Best Local Similarity 75.0%; Pred. No. 6.5e-25;
Matches 75; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
Qy 1 LKGI DESDSE DYVKEGLRAPLQSELDAR TKLTLELS DKIDELDAEIPKLEKNVYFK 60
Db 1 LKEI DESDSE DYIKEGLRAPLQSKLDAR KAKLSKLDSELS DKIDELDAEIAKLEKDVGFDP 60
Qy 61 LTDAEQTEQYLAAAEKDLADKKAELEKTEADLKAVHEPE 100
Db 61 NSDGEQAGQYLVAAEKDLDAKEAEELGNTGADLKKAVDEPE 100
RESULT 12
AAW14566
ID AAW14566 standard; protein; 185 AA.
XX
XX AAW14566;
XX
XX 17-OCT-2003 (revised)
XX
XX 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Ac94.
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
```

PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-Oct-2003 to standardise OS field)  
 XX  
 SQ Sequence 185 AA;  
 Query Match 68.7%; Score 339.5; DB 2; Length 185;  
 Best Local Similarity 74.3%; Pred. No. 8.7e-24;  
 Matches 75; Conservative 7; Mismatches 18; Indels 1; Gaps 1;  
 QY 1 LKGDSESDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAETPK-LEKNVEVF 59  
 DB 1 LKEDSESDYVKEGLRVPLQSELDVQAKLLEELSDKIDELDAETAKNKKQVDF 60  
 QY 60 KLTDAEQTYLAAAEKDLADKAELEKTEADLKAVHEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKAVNEPE 101  
 QY 60 KLTDAEQTYLAAAEKDLADKAELEKTEADLKAVHEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKAVNEPE 101  
 RESULT 13  
 ID AAW14571 standard; protein; 185 AA.  
 AC AAW14571;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Ac94c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 OS US6592876-B1.  
 PN 15-JUL-2003.  
 XX  
 PD 15-SEP-1995; 95US-00529055.  
 PF 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 DR Immunological composition for obtaining expression products used for  
 XX detecting the presence of Streptococcus pneumoniae or its strain,  
 PT

PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 46; 121pp; English.  
 PS  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 185 AA;  
 Query Match 68.7%; Score 339.5; DB 7; Length 185;  
 Best Local Similarity 74.3%; Pred. No. 8.7e-24;  
 Matches 75; Conservative 7; Mismatches 18; Indels 1; Gaps 1;  
 QY 1 LKGDSESDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAETPK-LEKNVEVF 59  
 DB 1 LKEDSESDYVKEGLRVPLQSELDVQAKLLEELSDKIDELDAETAKNKKQVDF 60  
 QY 60 KLTDAEQTYLAAAEKDLADKAELEKTEADLKAVHEPE 100  
 DB 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKAVNEPE 101  
 RESULT 14  
 ID AAW14571 standard; protein; 204 AA.  
 XX  
 AC AAW14571;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Ef1019.  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 1997-202002/18.  
 XX  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 PS Example 6; Fig 13; 296pp; English.  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-Oct-2003 to standardise OS field)  
 XX

CC surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 204 AA;  
Query Match 64.9%; Score 320.5; DB 2; Length 204;  
Best Local Similarity 68.2%; Pred. No. 5.9e-22;  
Matches 75; Conservative 3; Mismatches 11; Indels 21; Gaps 3;  
QY 1 LKGIDSDSDYVKEGLRAPLQSELDKRTKLTLELSKIDELDAEIPKLE----- 53  
DB 1 LKEIDSDSDYVKEGLRAPLQSELDKRTKLTLELSKIDELDAEIPKLE----- 60  
QY 54 --KNVE-YFKLTDAEQTEQYLAAAEKDLADKKAELEKTEADLKKAVNEPE 100  
DB 61 ENNVEDYFK-----EGLEKTTIAAKKAELEKTEADLKKAVNEPE 99

Search completed: June 18, 2005, 16:51:22  
Job time : 73.0731 secs

RESULT 15  
ABW02605  
ID ABW02605 standard; protein; 204 AA.  
XX  
AC ABW02605;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Ef1019c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
FA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
DR WPI; 2003-862841/80.  
XX  
XX  
PT Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 51; 121pp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspA) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-8  
Perfect score: 494  
Sequence: 1 LKGI DESDSEYVKEGLRAP.....KKALEKTEADLKKAHVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	100.0	100	4	US-09-147-875A-8
2	446.5	90.4	101	2	US-08-710-743-7
3	406	82.2	100	4	US-09-147-875A-2
4	399	80.8	550	4	US-09-583-110-4871
5	399	80.8	550	4	US-09-107-433-3858
6	396	80.2	100	4	US-09-147-875A-3
7	393	79.6	100	4	US-09-147-875A-5
8	392	79.4	98	4	US-09-147-875A-1
9	388.5	78.6	101	2	US-08-710-749-1
10	386	78.1	194	4	US-08-529-055-64
11	386	78.1	8991	4	US-08-714-741-32
12	381.5	77.2	101	2	US-08-710-749-2
13	376	76.1	100	4	US-09-147-875A-4
14	375.5	76.0	101	2	US-08-710-743-4
15	374.5	75.8	99	2	US-08-710-749-9
16	372	75.3	183	4	US-08-529-055-50
17	370	74.9	168	4	US-08-529-055-55
18	367	74.3	100	4	US-09-147-875A-6
19	365.5	74.0	101	2	US-08-710-743-3
20	356.5	72.2	101	2	US-08-710-749-5
21	351	71.1	166	4	US-08-529-055-48
22	346.5	70.1	101	4	US-09-147-875A-9
23	339.5	68.7	185	4	US-08-529-055-46
24	336	68.0	100	4	US-09-147-875A-7
25	329	66.6	102	2	US-08-710-743-8
26	323.5	65.5	101	2	US-08-710-749-6
27	320.5	64.9	99	2	US-08-710-749-10

28	320.5	64.9	99	4	US-09-147-875A-11	Sequence 11, Appl
29	320.5	64.9	204	4	US-08-529-055-51	Sequence 51, Appl
30	310.5	62.9	170	4	US-08-529-055-60	Sequence 60, Appl
31	310.5	62.9	181	4	US-08-529-055-42	Sequence 42, Appl
32	310.5	62.9	864	4	US-08-714-741-40	Sequence 40, Appl
33	309.5	62.7	99	2	US-08-710-743-11	Sequence 11, Appl
34	309.5	62.7	198	4	US-08-529-055-61	Sequence 61, Appl
35	309.5	62.7	619	1	US-08-465-746-2	Sequence 2, Appl
36	309.5	62.7	619	1	US-08-214-164-2	Sequence 2, Appl
37	309.5	62.7	619	2	US-08-467-852A-3	Sequence 3, Appl
38	309.5	62.7	619	2	US-08-246-636-2	Sequence 2, Appl
39	309.5	62.7	619	2	US-08-247-491A-3	Sequence 2, Appl
40	309.5	62.7	619	2	US-08-319-793-2	Sequence 2, Appl
41	309.5	62.7	619	2	US-08-468-985-2	Sequence 2, Appl
42	309.5	62.7	619	3	US-08-312-949-2	Sequence 2, Appl
43	309.5	62.7	648	1	US-08-072-070-2	Sequence 2, Appl
44	309.5	62.7	648	1	US-08-469-434-2	Sequence 2, Appl
45	309.5	62.7	648	1	US-08-214-222-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-147-875A-8  
; Sequence 8, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-8

Query Match 100.0%; Score 494; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.8e-38;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LKGI DESDSEYVKEGLRAPLQSELDKRTKLSTLEELSDKIDELDAEIPKLEKNVYFK 60  
Db 1 LKGI DESDSEYVKEGLRAPLQSELDKRTKLSTLEELSDKIDELDAEIPKLEKNVYFK 60  
Qy 61 LTDAEQTEQYLAALAAEKDLADKKALEKTEADLKKAHVHEPE 100  
Db 61 LTDAEQTEQYLAALAAEKDLADKKALEKTEADLKKAHVHEPE 100

RESULT 2  
US-08-710-749-7  
; Sequence 7, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/710,749  
;; FILING DATE: 20-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 101 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
US-08-710-749-7

Query Match 90.4%; Score 446.5; DB 2; Length 101;  
Best Local Similarity 93.1%; Pred. No. 4.1e-34;  
Matches 94; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 LKGI DESDSEYVKEGLRAPLQSEL-DAKRTKLTLELSKDIDELDAEIPKLEKNVEYF 59  
Db 1 LKGI DESDSEYVKEGLRAPLQSELDDAKQRTLTLELSKDIDELDAEIAKLEKNVEYF 60

Qy 60 KLTDAEQTEQYLA AAEKDLADKKA ELEKTEADLKKAVHEPE 100  
Db 61 KKTDAEQTEQYLA AAEKDLADKKA ELEKTEADLKKAVNEPE 101

RESULT 3  
US-09-147-875A-2  
; Sequence 2, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-2

Query Match 82.2%; Score 406; DB 4; Length 100;  
Best Local Similarity 85.0%; Pred. No. 2.2e-30;  
Matches 85; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKGI DESDSEYVKEGLRAPLQSELDAKRTKLTLELSKDIDELDAEIPKLEKNVEYF 60  
Db 1 LKGI DESDSEYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLA AAEKDLADKKA ELEKTEADLKKAVHEPE 100  
Db 61 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVHEPE 100

RESULT 4  
US-09-583-110-4871  
; Sequence 4871, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:

;; APPLICANT: Lynn Doucette-Stamm et al.  
;; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
;; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
;; FILE REFERENCE: PATH00-07A  
;; CURRENT APPLICATION NUMBER: US/09/583,110  
;; CURRENT FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/107,433  
;; PRIOR FILING DATE: 1998-06-30  
;; PRIOR APPLICATION NUMBER: US 60/085,131  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: US 60/051,553  
;; PRIOR FILING DATE: 1997-07-02  
;; NUMBER OF SEQ ID NOS: 5322  
;; SEQ ID NO 4871  
;; LENGTH: 550  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4871

Query Match 80.8%; Score 399; DB 4; Length 550;  
Best Local Similarity 84.0%; Pred. No. 6.8e-29;  
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKGI DESDSEYVKEGLRAPLQSELDAKRTKLTLELSKDIDELDAEIPKLEKNVEYF 60  
Db 144 LKGI DESDSEYVKEGLRAPLQSELDAKQAKLSKLELSKDIDELDAEIAKLEKNVEDFK 203  
Qy 61 LTDAEQTEQYLA AAEKDLADKKA ELEKTEADLKKAVHEPE 100  
Db 204 NSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 5  
US-09-107-433-3858  
; Sequence 3858, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:

;; APPLICANT: Lynn A Doucette-Stamm and David Bush  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
;; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
;; THERAPEUTICS

;; NUMBER OF SEQUENCES: 5206  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
;; STREET: 100 Beaver Street  
;; CITY: Waltham  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02354

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: CD-ROM ISO9660  
;; COMPUTER: <Unknown>  
;; OPERATING SYSTEM: <Unknown>  
;; SOFTWARE: <Unknown>  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/107,433  
;; FILING DATE: 30-Jun-1998  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/ 085131  
;; FILING DATE: May 12, 1998  
;; APPLICATION NUMBER: 60/051553  
;; FILING DATE: July 2, 1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Ariniello, Pamela Deneke  
;; REGISTRATION NUMBER: 40,489  
;; REFERENCE/DOCKET NUMBER: GTC-011  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (781) 893-5007  
;; TELEFAX: (781) 893-8277

;; INFORMATION FOR SEQ ID NO: 3858:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 550 amino acids  
;; TYPE: amino acid

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match      80.8%; Score 399; DB 4; Length 550;
Best Local Similarity 84.0%; Pred. No. 6.8e-29;
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||
Db 144 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 203
   |||||||

QY 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
   ::|||
Db 204 NSNGEAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 243
   ::|||

RESULT 6
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

Query Match      80.2%; Score 396; DB 4; Length 100;
Best Local Similarity 83.0%; Pred. No. 1.8e-29;
Matches 83; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||
Db 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||

QY 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
   ::|||
Db 61 NSNGEAEQYRAAAEDLAAKQAELEKTEADLKKAVHEPE 100
   ::|||

RESULT 7
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5
```

```
Query Match      79.6%; Score 393; DB 4; Length 100;
Best Local Similarity 83.0%; Pred. No. 3.4e-29;
Matches 83; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||
Db 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||

QY 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
   ::|||
Db 61 NSDGEQAGQYLAABEDLIKKAELKTEADLKKAVHEPE 100
   ::|||

RESULT 8
US-09-147-875A-1
; Sequence 1, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-1

Query Match      79.4%; Score 392; DB 4; Length 98;
Best Local Similarity 86.0%; Pred. No. 4.1e-29;
Matches 86; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

QY 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||
Db 1 LKGI DESSEDYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||||

QY 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
   ::|||
Db 61 NSDGEQA-QYLAABEDLIA-KKAELKTEADLKKAVHEPE 98
   ::|||

RESULT 9
US-08-710-749-1
; Sequence 1, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
```

```
;
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-1

Query Match 78.6%; Score 388.5; DB 2; Length 101;
Best Local Similarity 83.2%; Pred. No. 8.9e-29;
Matches 84; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKGIDESDSEYVKEGLRAPLQSEL-DAKRTKLTLEELSDKIDELDAEIPKLEKNVEYF 59
Db 1 LKEIDESDSEYVKEGLRAPLQSELDDAKQAKLSKLELSKIDELDAEIAKLEKNVEDF 60

Qy 60 KLTDAEQTEQYLAALAAEKDLADKKAELKTEADLKKAVHEP 100
Db 61 KNSGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVNEP 101

RESULT 10
US-08-529-055-64
; Sequence 64, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/529,055
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64

;
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-1

Query Match 78.6%; Score 388.5; DB 2; Length 101;
Best Local Similarity 83.2%; Pred. No. 8.9e-29;
Matches 84; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKGIDESDSEYVKEGLRAPLQSEL-DAKRTKLTLEELSDKIDELDAEIPKLEKNVEYF 59
Db 1 LKEIDESDSEYVKEGLRAPLQSELDDAKQAKLSKLELSKIDELDAEIAKLEKNVEDF 60

Qy 60 KLTDAEQTEQYLAALAAEKDLADKKAELKTEADLKKAVHEP 100
Db 61 KNSGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVNEP 101

RESULT 10
US-08-529-055-64
; Sequence 64, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/529,055
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64

;
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-1

Query Match 78.1%; Score 386; DB 4; Length 194;
Best Local Similarity 81.8%; Pred. No. 3.2e-28;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 1 LKEIDESDSEYVKEGLRAPLQSELDAKQAKLSKLELSKIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLAALAAEKDLADKKAELKTEADLKKAVHEP 99
Db 61 XSDGEQAGQYLAALAAEEDLIAKAELEQTEADLKKAVNEP 99

RESULT 11
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions Thereof,
; TITLE OF INVENTION: Expression Products Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/714,741
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match 78.1%; Score 386; DB 4; Length 8991;
Best Local Similarity 81.8%; Pred. No. 2.7e-26;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 7537 LKEIDESDSEYVKEGLRAPLQSELDAKQAKLSKLELSKIDELDAEIAKLEKNVEDFK 7596

Qy 61 LTDAEQTEQYLAALAAEKDLADKKAELKTEADLKKAVHEP 99
Db 61 XSDGEQAGQYLAALAAEEDLIAKAELEQTEADLKKAVNEP 99
```



```

RESULT 13
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100

```

RESULT 15

US-08-710-749-9  
 ; Sequence 9, Application US/08710749  
 ; Patent No. 5955089  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David B.  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Becker, Robert  
 ; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
 ; TITLE OF INVENTION: PROTEINS  
 ; NUMBER OF SEQUENCES: 28  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/710,749  
 ; FILING DATE: 20-SEP-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer, William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454312-2074  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 99 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: n/a  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: amino acid  
 US-08-710-749-9

Query Match 75.8%; Score 374.5; DB 2; Length 99;  
 Best Local Similarity 84.2%; Pred. No. 1.7e-27;  
 Matches 85; Conservative 5; Mismatches 8; Indels 3; Gaps 3;  
 Qy 1 LKGI DESDSEYVKEGLRAPLQSEL-DAKRTKLTLELSKIDELDAEIPKLEKNVEYF 59  
 Db 1 LKIDESEYVKEGLRAPLQSELDDAKQKLELSKIDELDAEIPKLEKNVEYF 60  
 Qy 60 KLTDAEQTEQYLAALAEKDLADKAELEKTEADLKAVHEPE 100  
 Db 61 KNSDGEQA-QYLAALAEEDLA-KKAELEKTEADLKAVNEPE 99

Search completed: June 18, 2005, 17:07:07  
 Job time : 18.9189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-8

Perfect score: 494

Sequence: 1 LKGI DESDSEYVKEGLRAP.....KKALEKTEADLKKAHVHEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10F\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	100.0	100	15	US-10-674-755-8
2	406	82.2	100	15	US-10-674-755-2
3	396	80.2	100	15	US-10-674-755-3
4	393	79.6	100	15	US-10-674-755-5
5	392	79.4	98	15	US-10-674-755-1
6	386	78.1	194	15	US-10-299-636-79
7	376	76.1	100	15	US-10-674-755-4
8	372	75.3	183	15	US-10-299-636-65
9	370	74.9	168	15	US-10-299-636-70
10	367	74.3	100	15	US-10-674-755-6
11	351	71.1	166	15	US-10-299-636-63
					Sequence 8, Appli
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 79, Appli
					Sequence 4, Appli
					Sequence 65, Appli
					Sequence 70, Appli
					Sequence 6, Appli

12	346.5	70.1	101	15	US-10-674-755-9	Sequence 9, Appli
13	339.5	68.7	185	15	US-10-299-636-61	Sequence 61, Appli
14	336	68.0	100	15	US-10-674-755-7	Sequence 7, Appli
15	320.5	64.9	99	15	US-10-674-755-11	Sequence 11, Appli
16	320.5	64.9	204	15	US-10-299-636-66	Sequence 66, Appli
17	310.5	62.9	170	15	US-10-299-636-75	Sequence 75, Appli
18	310.5	62.9	181	15	US-10-299-636-57	Sequence 57, Appli
19	310.5	62.9	643	15	US-10-299-636-95	Sequence 95, Appli
20	310.5	62.9	670	9	US-09-748-875-63	Sequence 63, Appli
21	310.5	62.9	670	10	US-09-298-523B-63	Sequence 61, Appli
22	310.5	62.9	690	9	US-09-748-875-61	Sequence 61, Appli
23	310.5	62.9	690	10	US-09-298-523B-61	Sequence 61, Appli
24	310.5	62.9	691	9	US-09-748-875-1	Sequence 1, Appli
25	310.5	62.9	691	10	US-09-298-523B-1	Sequence 1, Appli
26	310.5	62.9	701	9	US-09-748-875-62	Sequence 62, Appli
27	310.5	62.9	701	10	US-09-298-523B-62	Sequence 62, Appli
28	310.5	62.9	707	9	US-09-748-875-2	Sequence 2, Appli
29	310.5	62.9	707	10	US-09-298-523B-2	Sequence 2, Appli
30	310.5	62.9	711	9	US-09-748-875-3	Sequence 3, Appli
31	310.5	62.9	711	10	US-09-298-523B-3	Sequence 3, Appli
32	310.5	62.9	739	17	US-10-732-923-3294	Sequence 3294, Ap
33	310.5	62.9	929	9	US-09-748-875-60	Sequence 60, Appli
34	310.5	62.9	929	10	US-09-298-523B-60	Sequence 60, Appli
35	310.5	62.9	929	15	US-10-299-636-94	Sequence 94, Appli
36	309.5	62.7	198	15	US-10-299-636-76	Sequence 76, Appli
37	309.5	62.7	354	15	US-10-299-636-105	Sequence 105, Appli
38	309.5	62.7	588	15	US-10-299-636-96	Sequence 96, Appli
39	309.5	62.7	619	10	US-09-882-774-1	Sequence 1, Appli
40	309.5	62.7	619	15	US-10-282-122A-73702	Sequence 73702, A
41	309.5	62.7	619	16	US-10-414-532-72	Sequence 72, Appli
42	308	62.3	100	15	US-10-674-755-10	Sequence 10, Appli
43	308	62.3	100	15	US-10-674-755-12	Sequence 12, Appli
44	307.5	62.2	188	15	US-10-299-636-74	Sequence 74, Appli
45	303.5	61.4	99	15	US-10-674-755-16	Sequence 16, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-674-755-8  
; Sequence 8, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-8

Query Match 100.0%; Score 494; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 9.6e-35;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKGI DESDSEYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKVEYFK 60  
DB 1 LKGI DESDSEYVKEGLRAPLQSELDKAKTKLSTLEELSDKIDELDAEIPKLEKVEYFK 60

QY 61 LTDAEQTQYLAAEKDLADKKALEKTEADLKKAHVHEPE 100  
DB 61 LTDAEQTQYLAAEKDLADKKALEKTEADLKKAHVHEPE 100

#### RESULT 2

```
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match      82.2%; Score 406; DB 15; Length 100;
Best Local Similarity 85.0%; Pred. No. 3e-27;
Matches 85; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDKAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 1 LKIDESESDYVKEGLRAPLQSELDKAKQKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLAAAEKDLADKAELEKTEADLKAVHEPE 100
Db 61 NSNGEAEQYRAAEEDLAAKQAELEKTEADLKAVHEPE 100

RESULT 3
US-10-674-755-3
; Sequence 3, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-3

Query Match      80.2%; Score 396; DB 15; Length 100;
Best Local Similarity 83.0%; Pred. No. 2.1e-26;
Matches 83; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDKAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 1 LKIDESESDYVKEGLRAPLQSELDKAKQKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLAAAEKDLADKAELEKTEADLKAVHEPE 100
Db 61 NSNGEAEQYRAAEEDLAAKQAELEKTEADLKAVHEPE 100

RESULT 4
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
```

```
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match      79.6%; Score 393; DB 15; Length 100;
Best Local Similarity 83.0%; Pred. No. 3.9e-26;
Matches 83; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDKAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 1 LKIDESESDYVKEGLRAPLQSELDKAKQKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLAAAEKDLADKAELEKTEADLKAVHEPE 100
Db 61 NSDGEAQYLAFAEEDLIAKAELEKTEADLKAVHEPE 100

RESULT 5
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match      79.4%; Score 392; DB 15; Length 98;
Best Local Similarity 86.0%; Pred. No. 4.6e-26;
Matches 86; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

Qy 1 LKGIDESDSEYVKEGLRAPLQSELDKAKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 1 LKIDESESDYVKEGLRAPLQSELDKAKQKLSKLEELSDKIDELDAEIAKLEKNVEDFK 60

Qy 61 LTDAEQTEQYLAAAEKDLADKAELEKTEADLKAVHEPE 100
Db 61 NSDGEAQYLAFAEEDLIAKAELEKTEADLKAVHEPE 98

RESULT 6
US-10-299-636-79
; Sequence 79, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
```

```
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match      78.1%; Score 386; DB 15; Length 194;
Best Local Similarity 81.8%; Pred. No. 3.3e-25;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||
DB 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||

QY 61 LTDAEQTEQYLAARAEKDLADKAELEKTEADLKAVHEPE 99
   :|||
DB 61 XSDGEQAGQYLAARAEEDLIAKAELEKTEADLKAVNEP 99
   :|||

RESULT 7
US-10-674-755-4
; Sequence 4, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match      76.1%; Score 376; DB 15; Length 100;
Best Local Similarity 81.0%; Pred. No. 1.1e-24;
Matches 81; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||
DB 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||

QY 61 LTDAEQTEQYLAARAEKDLADKAELEKTEADLKAVHEPE 100
   :|||
DB 61 NSDGEQAGQYLAARAEEDLIAKAELEKTEADLKAVDEPE 100
   :|||

RESULT 8
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; CURRENT APPLICATION NUMBER: US/10299636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa at position 38 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa at position 42 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70
```

```
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-65

Query Match      75.3%; Score 372; DB 15; Length 183;
Best Local Similarity 79.0%; Pred. No. 4.8e-24;
Matches 79; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||
DB 1 LKIDSDSDYVYKEGLRAPLQSELDKARTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
   |||||

QY 61 LTDAEQTEQYLAARAEKDLADKAELEKTEADLKAVHEPE 100
   :|||
DB 61 NSDGEQAGQYLAARAEEDLIAKAELEKTEADLKAVDEPE 100
   :|||

RESULT 9
US-10-299-636-70
; Sequence 70, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa at position 38 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa at position 42 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70
```



; SOFTWARE: PatentIn Ver. 2.1

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-8  
Perfect score: 494  
Sequence: 1 LKIDSESDYVKEGLRAP.....KKAELEKTEADLKAVHEPE 100  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	309.5	62.7	619	2 A97887	surface protein ps
2	309.5	62.7	619	2 A41971	surface protein ps
3	119	24.1	744	2 P95013	pneumococcal surfa
4	112	22.7	1169	2 A64505	P115 homolog - Met
5	109	22.1	1319	2 A28313	glued protein - fr
6	108	21.9	1269	2 F84730	probable myosin he
7	104.5	21.2	1053	2 A41642	dynactin - chicken
8	101	20.4	629	2 T44607	hypothetical prote
9	100.5	20.3	886	2 H63378	conserved hypothet
10	100.5	20.3	1475	2 T33318	hypothetical prote
11	100	20.2	1179	2 F71190	probable chromosom
12	97	19.6	1620	2 S61535	nucleotide-binding
13	97	19.6	2116	2 A26655	myosin heavy chain
14	95	19.2	396	2 AE2137	hypothetical prote
15	95	19.2	1006	2 C70445	ATPase subunit of
16	94	19.0	879	2 C71083	conserved hypothet
17	94	19.0	1190	2 E84193	chromosome segrega
18	93	18.8	560	2 C71155	hypothetical prote
19	93	18.8	1312	2 T30845	probable DNA repai
20	92.5	18.7	408	2 S30283	protein M precursor
21	92.5	18.7	2139	2 T18296	myosin heavy chain
22	92.5	18.7	3450	2 T28563	hypothetical prote
23	92.5	18.7	3461	2 T26964	hypothetical prote
24	92	18.6	161	2 S48396	tropomyosin TPM2 -
25	91.5	18.5	436	2 S30284	M protein precursor
26	91.5	18.5	472	2 S43554	plasmidogen-bindin
27	91.5	18.5	1959	1 A33977	myosin heavy chain
28	91	18.4	387	2 S57834	fcra protein precu
29	91	18.4	388	2 A46173	Mrp4 protein - Str

30	91	18.4	1138	2 T24635	hypothetical prote
31	90.5	18.3	388	2 A49545	plasmidogen-bindin
32	90.5	18.3	415	2 S35760	fcra protein precu
33	90.5	18.3	482	2 T24518	hypothetical prote
34	90.5	18.3	522	2 G02533	occludin - human
35	90.5	18.3	532	2 S54871	M protein - Strept
36	90.5	18.3	896	2 A43074	epidermal growth f
37	90	18.2	405	2 A33939	Fc gamma (IgG) rec
38	90	18.2	577	1 A41289	moesin - human
39	90	18.2	577	1 S39804	moesin - pig
40	90	18.2	880	2 F75103	conserved hypothet
41	90	18.2	1310	2 T34063	chromosome segrega
42	89	18.0	230	2 I40287	outer surface prot
43	89	18.0	564	2 A60115	M protein precursor
44	89	18.0	1164	2 T24806	hypothetical prote
45	89	18.0	1186	2 G69708	chromosome segrega

ALIGNMENTS

RESULT 1

A97887 surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID: 1;  
C:Genetics:  
A:Gene: pspA

Query Match 62.7%; Score 309.5; DB 2; Length 619;  
Best Local Similarity 65.5%; Pred. No. 1.5e-14;  
Matches 72; Conservative 5; Mismatches 12; Indels 21; Gaps 3;  
QY 1 LKIDSESDYVKEGLRAPLOSELDARFKLSTLEELSDKIDELDAEIPKLE----- 53  
DB 223 LKIDSESDYVKEGLRAPLOSELDARFKLSTLEELSDKIDELDAEIPKLE----- 53  
QY 54 --KNVE-YFKLTDAFQTQYLAAAEKDLADKKAELEKTEADLKAVHEPE 100  
DB 283 ENNVEDYFK-----EGLEKTAACKAELEKTEADLKAVHEPE 321

RESULT 2

A41971 surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:g153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991

A:Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability

A:Reference number: A60282; MUID:91169598; PMID:2004810

A:Accession: A60282

A:Molecule type: protein

A:Residues: 32-76 <TAL>

A:Experimental source: strain JY2008

C:Genetics:

A:Gene: pspA

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-619/Product: surface protein pspA #status predicted <MAT>

F:411-430/Domain: cpl repeat homology <CP01>

F:431-450/Domain: cpl repeat homology <CP02>

F:451-470/Domain: cpl repeat homology <CP03>

F:471-490/Domain: cpl repeat homology <CP04>

F:491-510/Domain: cpl repeat homology <CP05>

F:511-530/Domain: cpl repeat homology <CP06>

F:531-550/Domain: cpl repeat homology <CP07>

F:551-570/Domain: cpl repeat homology <CP08>

F:571-591/Domain: cpl repeat homology <CP09>

F:592-611/Domain: cpl repeat homology <CP10>

Query Match 62.7%; Score 309.5; DB 2; Length 619;

Best Local Similarity 65.5%; Pred. No. 1.5e-14;

Matches 72; Conservative 5; Mismatches 12; Indels 21; Gaps 3;

Qy 1 LKGI DESDSDYVKEGLRAPLQSELDKRTKLTLELSKDIDELDAEIPKLE----- 53

Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKDIDELDAEIAKLEDLQKAAE 282

Qy 54 --KNVE-YFKLTDAEQTEQYLAARADKADKAELEKTEADLKKAVHEPE 100

Db 283 ENNVEDYFK-----EGLEKTTAAKKAELEKTEADLKKAVNEPE 321

RESULT 3

F95013

pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C:Accession: F95013

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple, nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: F95013

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-744 <KUR>

A:Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0117

Query Match 24.1%; Score 119; DB 2; Length 744;

Best Local Similarity 31.9%; Pred. No. 0.37;

Matches 38; Conservative 16; Mismatches 43; Indels 22; Gaps 3;

Qy 1 LKGI DESDSDYVKEGLRAPLQSELDKRTKLTSTI-----EELSD 40

Db 323 LGGADPEDDTAALQNKLAAR-KAELAKKQTELEKILSLDPGKTDQLDKEAEAEALDK 381

Qy 41 KIDELDAEIPKLEKNVEYFK-LTDAEQTEQYLAARADKADKAELEKTEADLKKAVHE 98

Db 382 KADELQNVADLEKEISNLEILLGGADSEDDTAALQNKLATYKAELEKTIQKELDAALNE 440

RESULT 4

A64505

P115 homolog - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Jun-2000

C:Accession: A64505

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.F.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: A64505

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1169 <BUL>

A:Cross-references: GB:U67604; GB:L77117; NID:gl592224; PID:gl500543; TIGR:MJ1643

C:Genetics:

A:Map position: FOR1623481-1626990

C:Superfamily: chromosome segregation protein SMC1

Query Match 22.7%; Score 112; DB 2; Length 1169;

Best Local Similarity 29.5%; Pred. No. 1.8;

Matches 33; Conservative 24; Mismatches 41; Indels 14; Gaps 2;

Qy 1 LKGI DESDSDYVKEGLRAPLQSELD-----AKRTKLTLELSKDIDELDAEIPKLEKN 55

Db 799 LKRWNEIEGELKILEKAKLKNKIDGLTLVKELLPKIEELNKKVSELINKKVVILEKN 858

Qy 56 VEYFK-----LTDAEQTEQYLAARADKADKAELEKTEADLKKAVHE 98

Db 859 ISPKYSEIKNLSILEKKRYEELAKNKLKELTEKKEQLEKIEITLERRE 910

RESULT 5

A28313

glued protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004

C:Accession: A28313

R:Swaroop, A.; Swaroop, M.; Garen, A.

Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987

A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ge

A:Reference number: A28313; MUID:87317680; PMID:2819881

A:Accession: A28313

A:Molecule type: DNA; mRNA

A:Residues: 1-1319 <SWA>

A:Cross-references: UNIPROT:P13496

A:Note: the authors' translation is inconsistent with the nucleotide sequence in the reg

C:Genetics:

A:Gene: FlyBase:Gl

A:Cross-references: FlyBase:FBgn0001108

A:Introns: 18/2; 479/3

C:Keywords: cytoskeleton; glycoprotein

F:397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 22.1%; Score 109; DB 2; Length 1319;

Best Local Similarity 33.7%; Pred. No. 3.4;

Matches 35; Conservative 20; Mismatches 35; Indels 14; Gaps 4;

Qy 1 LKGI DESDSDYVKEGLRAPLQSELDKRTKLS----TLELSKDIDELDAEIPKLEKNV 56

Db 429 LRDLISAHDKHDIQK-----LSKELEMKRSEVTELETKKLSAKIDELAEIVADLQEQV 482

Qy 57 EYFKLTDAEQTEQYLAARADKADKAELEKTEADLK--KAVHE 98

Db 483 D--AALGAEEVMEQLAEKQMELEDKVKLLEEEIAQLEAELEVHE 524

RESULT 6

F84730

probable myosin heavy chain [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: F84730

R.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: F84730  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1269 <STO>  
A;Cross-references: GB:AE002093; NID:g6598483; PIDN:AAC69932.2; GSPDB:GN00139  
A;Gene: At2g32240  
A;Map position: 2

Query Match 21.9%; Score 108; DB 2; Length 1269;  
Best Local Similarity 32.7%; Pred. No. 3.8;  
Matches 34; Conservative 21; Mismatches 35; Indels 14; Gaps 3;

QY 3 GIDSDSDYVYKGLRAPLOSLDRAKTKLSTLE-----ELSKIDELDAEIPK 51  
DB 658 GETEADSKGYL--GOVAELSTLEAFQVKSLSLEAALNIATENEKELTENLNAVTSKKK 715

QY 52 LKNNVEFKLTDAEQTEQYLAAEKDLADKKALEKTEADLKKA 95

DB 716 LEATVDSEYKISE--SENLLSETRNELNVQTKLESINDLKAA 758

## RESULT 7

A41642  
C;Species: Gallus gallus (chicken)  
C;Date: 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change 09-Jul-2004  
C;Accession: A41642  
R; Gill, S.R.; Schroer, T.A.; Szilak, I.; Steuer, B.R.; Sheetz, M.P.; Cleveland, D.W.  
J. Cell Biol 115, 1639-1650, 1991

A;Title: Dynactin, a conserved, ubiquitously expressed component of an activator of vesic  
A;Reference number: A41642; MUID:92098576; PMID:1836789

A;Accession: A41642

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1053 <GIL>

A;Cross-references: UNIPROT:P35458; GB:X62773; NID:g63920; PID:g63921

C;Keywords: cytoskeleton

Query Match 21.2%; Score 104.5; DB 2; Length 1053;  
Best Local Similarity 29.0%; Pred. No. 5.5;  
Matches 31; Conservative 24; Mismatches 29; Indels 23; Gaps 5;

QY 1 LKIDSDSDYVYKGLRAPLOSLDRAKTKLSTLESLDKIDELDAEIPKKNVEYFK 60  
DB 199 MRDLASAEKQEHVK-----LQKMEKKNTELESIRQREKLQE---EVKQAEKTVDELK 249

QY 61 LTDAEQTEQYLAAEK-----DLADKKALEKTEADLKKAHVE 98

DB 250 ----EQVDAALGAEMVETLTERNLDLSEKVRLETGVGL-EMANE 291

## RESULT 8

A44607  
C;Species: Halobacterium salinarum  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T44607

R; Ruepp, A.; Wanner, G.; Soppa, J.

Arch. Microbiol. 169, 1-9, 1998

A;Title: A 71-kDa protein from Halobacterium salinarum belongs to a ubiquitous P-loop A  
A;Reference number: Z22810; MUID:98060711; PMID:9396829

A;Accession: T44607

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-629 <RUE>

A;Cross-references: UNIPROT:O07116; EMBL:Y13615; PIDN:CAA73936.1

A;Note: the source is designated as Halobacterium salinarum  
C;Genetics:  
A;Note: hp71  
C;Function:  
A;Description: might be involved in cytoskeleton formation and/or chromosome partitioni

Query Match 20.4%; Score 101; DB 2; Length 629;  
Best Local Similarity 31.5%; Pred. No. 5.7;  
Matches 28; Conservative 22; Mismatches 37; Indels 2; Gaps 2;

QY 6 ESDSDYVYKGLRAPLOSLDRAKTKLSTLESLDKIDELDAEIPKKNVEYFKLTD 65  
DB 367 KAELEDEIR--LRVDIQEDQHEVRSIEATIEELQAEIQREAEYEAAGAGESHS-AELK 424

QY 66 QTEQYLAAAEKDLADKKALEKTEADLK 94

DB 425 TIQKIGSTETKLDRAQAELERAEALQK 453

## RESULT 9

H69378  
C;Species: Archaeoglobus fulgidus  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: H69378

R; Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso;

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Usterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB902

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 20.3%; Score 100.5; DB 2; Length 886;

Best Local Similarity 26.2%; Pred. No. 8.8;

Matches 34; Conservative 24; Mismatches 37; Indels 35; Gaps 4;

QY 1 LKIDSDSDYVYKGLRAPLOSLDRAKTKLSTLESLDKIDELDAEIPKKN----- 55  
DB 303 LRDVEKREG-DLTREA--AGIQALKAEDNSKLEIITKRIEELERELERFEKSHRLLE 359

QY 56 -----VF--YFKLTDAEQTEQYLAAAEKDLADKKALEK 88  
DB 360 TLKPKMDRMQGIKAKLEKNLTPDKVKMYDLSKAKEBEITEKLLKLIAKKSSLKTR 419

QY 89 EADLKKAHVE 98

DB 420 GAQLKKAHVE 429

## RESULT 10

T33318  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T33318

R; Geisel, C.; Bradshaw, H.

submitted to the EMBL Data Library, May 1998

A;Description: The sequence of C. elegans cosmid ZK1055.

A;Reference number: Z21321

A;Accession: T33318

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1475 <GEI>

A;Cross-references: UNIPROT:O76447; EMBL:AF068721; PIDN:AAC19259.1; GSPDB:GN00023; CESP

A;Experimental source: strain Bristol N2; clone ZK1055



A;Reference number: AB1807; MUID:21595285; PMID:11759840

Query Match 19.2%; Score 95; DB 2; Length 396;

**RESULT 15**

Query Match 19.2%; Score 95; DB 2; Length 1006;

Search completed: June 18, 2005, 17:03:55

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:34:50 ; Search time 60.961 Seconds  
(without alignments)  
840.012 Million cell updates/sec

Title: US-10-674-755-8  
Perfect score: 494  
Sequence: 1 LKGIQSDSDYVKEGLRAP.....KKALEKTEADLKAVHEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473	95.7	237	Q9L592	Q9L592 streptococc
2	473	95.7	395	Q9LAY9	Q9LAY9 streptococc
3	447	90.5	207	Q8GNS9	Q8GNS9 streptococc
4	399	80.8	222	Q9L577	Q9L577 streptococc
5	399	80.8	262	Q9L576	Q9L576 streptococc
6	399	80.8	415	Q9LAY7	Q9LAY7 streptococc
7	392	79.4	416	Q9LAY8	Q9LAY8 streptococc
8	390	78.9	225	Q9L591	Q9L591 streptococc
9	389	78.7	406	Q9LAZ0	Q9LAZ0 streptococc
10	388	78.5	194	Q9L5B5	Q9L5B5 streptococc
11	388	78.5	218	Q6UEB2	Q6UEB2 streptococc
12	388	78.5	233	Q9L568	Q9L568 streptococc
13	388	78.5	243	Q9L569	Q9L569 streptococc
14	388	78.5	243	Q9L564	Q9L564 streptococc
15	388	78.5	243	Q9L567	Q9L567 streptococc
16	388	78.5	244	Q9L565	Q9L565 streptococc
17	388	78.5	247	Q9L566	Q9L566 streptococc
18	388	78.5	249	Q9L570	Q9L570 streptococc
19	388	78.5	254	Q9L563	Q9L563 streptococc
20	388	78.5	340	Q8KQK5	Q8KQK5 streptococc
21	388	78.5	401	Q9LAZ2	Q9LAZ2 streptococc
22	386	78.1	394	Q9LAY6	Q9LAY6 streptococc
23	386	78.1	395	Q9LAZ1	Q9LAZ1 streptococc
24	383	77.5	246	Q9L578	Q9L578 streptococc
25	378	76.5	255	Q9L581	Q9L581 streptococc
26	378	76.5	255	Q9L5B6	Q9L5B6 streptococc
27	358	72.5	393	Q9LAZ3	Q9LAZ3 streptococc
28	320.5	64.9	417	Q9LAZ3	Q9LAZ3 streptococc
29	310.5	62.9	739	Q9RQT4	Q9RQT4 streptococc
30	310.5	62.9	820	Q9RQT1	Q9RQT1 streptococc
31	310.5	62.9	929	Q9KK19	Q9KK19 streptococc

32	310.5	62.9	929	2	Q9ZAYS	Q9ZAYS streptococc
33	309.5	62.7	619	2	Q54972	Q54972 streptococc
34	309.5	62.7	619	2	Q8DR10	Q8DR10 streptococc
35	302.5	61.2	415	2	Q9LAY1	Q9LAY1 streptococc
36	298.5	60.4	99	2	Q8KQK4	Q8KQK4 streptococc
37	298.5	60.4	437	2	Q9LAY4	Q9LAY4 streptococc
38	292.5	59.2	426	2	Q9LAY5	Q9LAY5 streptococc
39	291.5	59.0	249	2	Q9L575	Q9L575 streptococc
40	288.5	58.4	224	2	Q8GNS8	Q8GNS8 streptococc
41	280	56.7	869	2	Q9KK27	Q9KK27 streptococc
42	279.5	56.6	395	2	Q9LAY2	Q9LAY2 streptococc
43	279.5	56.6	408	2	Q9LAY0	Q9LAY0 streptococc
44	155	31.4	479	2	Q9LAX2	Q9LAX2 streptococc
45	155	31.4	480	2	Q9LAX3	Q9LAX3 streptococc

ALIGNMENTS

RESULT 1

Q9L592 PRELIMINARY; PRT; 237 AA.  
AC Q9L592;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP194;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones".  
RT J. Clin. Microbiol. 38:3663-3669(2000).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=SP194;  
RC STRAIN=SP194;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF254257; AAF68092.1; -  
FT NON\_TER 1  
FT NON\_TER 237  
SQ SEQUENCE 237 AA; 26638 MW; 345B91580B6D372 CRC64;

Query Match 95.7%; Score 473; DB 2; Length 237;

Best Local Similarity 97.0%; Pred. No. 4.8e-23;  
Matches 97; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LKGIQSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIKPKKNVEYFK 60  
|||||

Db 47 LKGIQSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIKPKKNVEYFK 106  
|||||

QY 61 LTDABQTEQYLAABEKDLADKKALEKTEADLKAVHEPE 100  
|||||

Db 107 KTDABQTEQYLAABEKDLADKKALEKTEADLKAVHEPE 146  
|||||

RESULT 2

Q9LAY9 PRELIMINARY; PRT; 395 AA.  
AC Q9LAY9;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;

```
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL1;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/JAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071806; AAF27702.1; -.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44287 MW; 9AC2301BABB37F90 CRC64;

Query Match 95.7%; Score 473; DB 2; Length 395;
Best Local Similarity 97.0%; Pred. No. 7, 6e-23;
Matches 97; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 213 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 272
Qy 61 LTDAEQTEQYLA AAEKDLADKAELEKTEADLKKA VHEPE 100
Db 273 KTD AEQTEQYLA AAEKDLADKAELEKTEADLKKA VNEPE 312

RESULT 3
Q8GNS9 PRELIMINARY; PRT; 207 AA.
AC Q8GNS9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP356;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicunzio G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490266; AAN37734.1; -.
FT NON_TER 1
FT NON_TER 207
SQ SEQUENCE 207 AA; 23135 MW; 809021CS9D9D71A43 CRC64;

Query Match 90.5%; Score 447; DB 2; Length 207;
Best Local Similarity 91.0%; Pred. No. 2e-21;
Matches 91; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 15 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 74
Qy 61 LTDAEQTEQYLA AAEKDLADKAELEKTEADLKKA VHEPE 100
Db 75 QTN AEQTEQYLA AAEKDLVAKAELEKTEADLKKA VNDPE 114

RESULT 4
Q9L577 PRELIMINARY; PRT; 222 AA.
ID Q9L577
```

```
AC Q9L577;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255550; AAF68103.1; -.
FT NON_TER 1
FT NON_TER 222
SQ SEQUENCE 222 AA; 24558 MW; 6D7EB7842FE9F2A6 CRC64;

Query Match 80.8%; Score 399; DB 2; Length 222;
Best Local Similarity 84.0%; Pred. No. 2.5e-18;
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db 25 LKGI DESDSE DYVKEGLRAPLQSELD AKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 84
Qy 61 LTDAEQTEQYLA AAEKDLADKAELEKTEADLKKA VHEPE 100
Db 85 NSNGEAEQYRA AAEEDLA AKQAELEKTEADLKKA VNEPE 124

RESULT 5
Q9L576 PRELIMINARY; PRT; 262 AA.
ID Q9L576
AC Q9L576;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255551; AAF68104.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
```



```
FT NON TER 262 262
SQ SEQUENCE 262 AA; 29012 MW; 32C76909946A584 CRC64;

Query Match
Best Local Similarity 80.8%; Score 399; DB 2; Length 262;
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKGDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEYFK 60
Db 65 LKIDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEYFK 124

QY 61 LTDAEQTEQYLAARAKDLADKKAELKTEADLKKAVHEPE 100
Db 125 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 164

RESULT 6
Q9LAY7
ID Q9LAY7 PRELIMINARY; PRT; 415 AA.
AC Q9LAY7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG6692;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071808; AAF27704.1; -
FT NON TER 415 415
SQ SEQUENCE 415 AA; 45593 MW; 41375ACBFA10FA46 CRC64;

Query Match
Best Local Similarity 80.8%; Score 399; DB 2; Length 415;
Matches 84; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKGDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEYFK 60
Db 229 LKIDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEDFK 288

QY 61 LTDAEQTEQYLAARAKDLADKKAELKTEADLKKAVHEPE 100
Db 289 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 328

RESULT 7
Q9LAY8
ID Q9LAY8 PRELIMINARY; PRT; 416 AA.
AC Q9LAY8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8838;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
```

```
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071807; AAF27703.1; -
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 416 416
SQ SEQUENCE 416 AA; 45987 MW; 990C8858BC6B12C7 CRC64;

Query Match
Best Local Similarity 79.4%; Score 392; DB 2; Length 416;
Matches 83; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKGDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEYFK 60
Db 229 LKIDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEDFK 288

QY 61 LTDAEQTEQYLAARAKDLADKKAELKTEADLKKAVHEPE 100
Db 289 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 328

RESULT 8
Q9L591
ID Q9L591 PRELIMINARY; PRT; 225 AA.
AC Q9L591;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP199;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254258; AAF68093.1; -
DR InterPro; IPR009082; His_kin_homodim.
FT NON TER 1 1
FT NON TER 225 225
SQ SEQUENCE 225 AA; 24835 MW; F878A7618B72A692 CRC64;

Query Match
Best Local Similarity 78.9%; Score 390; DB 2; Length 225;
Matches 82; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 LKGDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEYFK 60
Db 34 LKIDSESDSDYVKEGLRAPLQSELDKRTKLTSTLEELSDKIDELDAETPKLEKNVEDFK 93

QY 61 LTDAEQTEQYLAARAKDLADKKAELKTEADLKKAVHEPE 100
Db 94 NSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 133

RESULT 9
Q9LAZ0
ID Q9LAZ0 PRELIMINARY; PRT; 406 AA.
AC Q9LAZ0;
```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL6A;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071805; AAF27701.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 406
SQ SEQUENCE 406 AA; 45406 MW; 5B1E20CA4D06F052 CRC64;

Query Match 78.7%; Score 389; DB 2; Length 406;
Best Local Similarity 82.0%; Pred. No. 1.9e-17;
Matches 82; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 60
Db 213 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 60

Qy 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
Db 273 NSDGEQAEQYLAAAEEDLIANKAELEQTEADLKKAVNEPE 312

RESULT 10
Q9LSB5 PRELIMINARY; PRT; 194 AA.
AC Q9LSB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253407; AAF67355.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON TER 194
FT NON TER 194
SQ SEQUENCE 194 AA; 21116 MW; E68189FCA2B244F8 CRC64;

Query Match 78.5%; Score 388; DB 2; Length 194;
Best Local Similarity 81.0%; Pred. No. 1.1e-17;

```

```

Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 60
Db 55 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 114

Qy 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
Db 115 NSDGEQAEQYLVAALKDLDKKAELKTEADLKKAVDEPE 154

RESULT 11
Q6UEB2 PRELIMINARY; PRT; 218 AA.
AC Q6UEB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URSP2;
RX PubMed=14977961; DOI=10.1128/IAI.72.3.1548-1556.2004;
RA Wells J., Gigliotti F., Simpson-Haidaris P.J., Haidaris C.G.;
RT "Epitope mapping of a protective monoclonal antibody against
Pneumocystis carinii with shared reactivity to Streptococcus
pneumoniae surface antigen PspA.";
RL Infect. Immun. 72:1548-1556(2004).
DR EMBL; AY371665; AAR20918.1; -.
DR InterPro; IPR009082; His_kin_homodim.
FT NON TER 1
FT NON TER 218
SQ SEQUENCE 218 AA; 22926 MW; 8F9F27EDE0A08D72 CRC64;

Query Match 78.5%; Score 388; DB 2; Length 218;
Best Local Similarity 81.0%; Pred. No. 1.2e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 60
Db 27 LKGI DESDSEDYVKEGLRAPLQSELDAKRTKLTLELSKIDELDAEIPKLEKNVEYFK 86

Qy 61 LTDAEQTEQYLAAAEKDLADKKAELKTEADLKKAVHEPE 100
Db 87 NSDGEQAEQYLVAALKDLDKKAELKTEADLKKAVDEPE 126

RESULT 12
Q9LS68 PRELIMINARY; PRT; 233 AA.
AC Q9LS68;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";

```

```
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=39;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255902; AAF70092.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 233
SQ SEQUENCE 233 AA; 24514 MW; D5C494019C45BPB2 CRC64;

Query Match 78.5%; Score 388; DB 2; Length 233;
Best Local Similarity 81.0%; Pred. No. 1.3e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKGDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 LKIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 87
QY 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 NSDGEAQEQLVAAKDLDAKKALENTADLKKAVDEPE 127

RESULT 13
Q9L569 PRELIMINARY; PRT; 236 AA.
AC Q9L569;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=177;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255901; AAF70091.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 236
SQ SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;

Query Match 78.5%; Score 388; DB 2; Length 236;
Best Local Similarity 81.0%; Pred. No. 1.3e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKGDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 LKIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 108
QY 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
109 NSDGEAQEQLVAAKDLDAKKALENTADLKKAVDEPE 148

RESULT 14
Q9L564 PRELIMINARY; PRT; 243 AA.
AC Q9L564;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=152;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255906; AAF70096.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243
SQ SEQUENCE 243 AA; 26145 MW; 28D15207554137CB CRC64;

Query Match 78.5%; Score 388; DB 2; Length 243;
Best Local Similarity 81.0%; Pred. No. 1.4e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKGDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 LKIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 133
QY 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 NSDGEAQEQLVAAKDLDAKKALENTADLKKAVDEPE 173

RESULT 15
Q9L567 PRELIMINARY; PRT; 243 AA.
AC Q9L567;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255901; AAF70091.1; -.
FT InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 236
SQ SEQUENCE 236 AA; 25067 MW; AC09706F048018A5 CRC64;

Query Match 78.5%; Score 388; DB 2; Length 236;
Best Local Similarity 81.0%; Pred. No. 1.3e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKGDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 LKIDSDSDYVKEGLRAPLQSELDKRTKLTLEELSDKIDELDAEIPKLEKNVEYFK 108
QY 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
109 NSDGEAQEQLVAAKDLDAKKALENTADLKKAVDEPE 148
```

```

DR EMBL; AF255903; AAF70093.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR00533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 243 243
SQ SEQUENCE 243 AA; 25700 MW; D64F521365723B57 CRC64;

Query Match      78.5%; Score 388; DB 2; Length 243;
Best Local Similarity 81.0%; Pred. NO. 1.4e-17;
Matches 81; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LKGIDSESDYVKEGLRPIQSELDKRTKLTLEELSDKIDLDLAEIPKLEKNVEYFK 60
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 LKEIDSESDYIKEGLRPIQSKLDKAKKAKLSELSKIDLDLAEIAKLEKVEDFK 109

Qy 61 LTDAEQTEQYLAAAEKDLADKKALEKTEADLKXAVHEPE 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 NSDGEQAEQYLVAAKKDLDAKKALENTADLKXAVDEPE 149

```

Search completed: June 18, 2005, 17:01:36  
Job time : 60.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.8038 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-9  
Perfect score: 496  
Sequence: 1 LKEIDSESDYVKEGLRVP.....KKAELEKTEADLKXAVHEPE 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489	98.6	185	2	AAW14566 Streptococ
2	489	98.6	185	7	ABW02600 Ac94C pne
3	489	98.6	8991	6	ABU08487 S. pneumo
4	386.5	77.9	183	2	AAW14570 Streptococ
5	386.5	77.9	183	7	ABW02604 Streptococ
6	385.5	77.7	550	8	ADK48356 Streptococ
7	385.5	77.7	550	8	ADR95223 Novel S.
8	383.5	77.3	194	2	AAW14584 Streptococ
9	383.5	77.3	194	7	ABW02618 Dbl16ac pn
10	376.5	75.9	168	7	ABW02609 L81905c p
11	363	73.2	167	2	AAW14575 Streptococ
12	354.5	71.5	166	2	AAW14568 Streptococ
13	354.5	71.5	166	7	ABW02602 Bg8743c p
14	329	66.3	204	2	AAW14571 Streptococ
15	329	66.3	204	7	ABW02605 Efl1019c p
16	313	63.1	198	7	ABW02615 Rx1c pneu
17	313	63.1	315	2	AAW04375 Streptococ
18	313	63.1	619	2	AAW04377 Streptococ
19	313	63.1	619	2	AAW04379 Streptococ
20	313	63.1	619	2	AAW04381 Streptococ
21	313	63.1	619	2	AAW04383 Streptococ
22	313	63.1	619	5	AAW04385 Streptococ
23	313	63.1	619	6	AAW04387 Streptococ
24	313	63.1	619	8	AAW04389 Streptococ
25	313	63.1	648	2	AAW04391 Streptococ

26	313	63.1	648	2	AAW62274 Streptococ
27	313	63.1	648	2	AAW41837 Streptococ
28	313	63.1	648	2	AAW87879 A. pneumoc
29	313	63.1	653	2	AAW92456 S. pneumo
30	313	63.1	684	2	AAW73912 Streptococ
31	312	62.9	198	2	AAW14581 Streptococ
32	311	62.7	170	7	ABW02614 Rct135c p
33	311	62.7	181	7	ABW02596 0922134c
34	311	62.7	865	6	ABU08489 S. pneumo
35	311	62.7	929	2	AAW14593 Streptococ
36	311	62.7	929	2	AAW43384 S. pneumo
37	308	62.1	188	2	AAW14580 Streptococ
38	308	62.1	188	7	ABW02613 Rct129c p
39	304	61.3	195	2	AAW14591 Streptococ
40	304	61.3	195	7	ABW02625 Wu2c pneu
41	299	60.3	588	6	ABU08491 Coiled co
42	299	60.3	589	2	AAW43392 Streptococ
43	298	60.1	206	2	AAW14574 Streptococ
44	298	60.1	206	7	ABW02608 Dbl15c pne
45	297	59.9	204	2	AAW14578 Streptococ

## ALIGNMENTS

RESULT 1  
AAW14566  
ID AAW14566 standard; protein; 185 AA.  
XX  
AC AAW14566;  
XX  
DT 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
DE Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
OS Streptococcus pneumoniae; strain Ac94.

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac94.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine

```

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 185 AA;

Query Match          98.6%; Score 489; DB 2; Length 185;
Best Local Similarity 99.0%; Pred. No. 4.5e-41;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYVKEGLRVPLQSELDVQKALLKLELSKIDELDAIAKNLKKVDVF 60
   |||||
DB 1 LKEIDESDSEYVKEGLRVPLQSELDVQKALLKLELSKIDELDAIAKNLKKVDVF 60
   |||||

QY 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
   |||||
DB 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 101
   |||||

RESULT 2
ABW02600
ID ABW02600 standard; protein; 185 AA.
XX
AC ABW02600;
XX
DT 12-FEB-2004 (first entry)
XX
DE Ac94c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
DR WPI; 2003-862841/80.
XX
Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
Example 6; SEQ ID NO 46; 121pp; English.
XX
The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Ac94c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 185 AA;

Query Match          98.6%; Score 489; DB 7; Length 185;
Best Local Similarity 99.0%; Pred. No. 4.5e-41;

Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYVKEGLRVPLQSELDVQKALLKLELSKIDELDAIAKNLKKVDVF 60
   |||||
DB 4382 LKEIDESDSEYVKEGLRVPLQSELDVQKALLKLELSKIDELDAIAKNLKKVDVF 4441
   |||||

```

QY 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101  
 DB 4442 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 4482

## RESULT 4

AAW14570  
 ID AAW14570 standard; protein; 183 AA.

XX AC AAW14570;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE XX  
 KW PepA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Bg9739.

OS WO9709994-A1.  
 PN 20-MAR-1997.  
 PD 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 PR (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 DR Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg9739.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 183 AA;  
 SQ

Query Match 77.9%; Score 386.5; DB 2; Length 183;  
 Best Local Similarity 81.2%; Pred. No. 9.2e-31;  
 Matches 82; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 LKEIDESDESDYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNLKKVDPEF 60  
 DB 1 LKEIDESDESDYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNLKKVDPEF 59

QY 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101  
 DB 60 QNSDGEQAGQYLAAGADLIATAKAELEKAEADLKKAVDEPE 100

## RESULT 5

ABW02604  
 ID ABW02604 standard; protein; 183 AA.

XX ABW02604;  
 AC 12-FEB-2004 (first entry)

DE Bg9739c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX US6592876-B1.  
 PN 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 50; 121pp; English.

XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Bg9739c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention

XX Sequence 183 AA;  
 SQ

Query Match 77.9%; Score 386.5; DB 7; Length 183;  
 Best Local Similarity 81.2%; Pred. No. 9.2e-31;  
 Matches 82; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 LKEIDESDESDYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNLKKVDPEF 60  
 DB 1 LKEIDESDESDYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAETAKNLKKVDPEF 59

QY 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101  
 DB 60 QNSDGEQAGQYLAAGADLIATAKAELEKAEADLKKAVDEPE 100

## RESULT 6

ADK48356  
 ID ADK48356 standard; protein; 550 AA.

XX AC ADK48356;  
 XX 20-MAY-2004 (first entry)

DT Streptococcus pneumoniae protein, Seq ID No 4871.  
 DE

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
KW Streptococcus pneumoniae.  
OS US6699703-B1.  
XX 02-MAR-2004.  
XX 26-MAY-2000; 2000US-005683110.  
XX 02-JUL-1997; 97US-0051553P.  
XX 12-MAY-1998; 98US-0085131P.  
XX 30-JUN-1998; 98US-00107433.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;  
PI WPI; 2004-212399/20.  
DR N-ESDB; ADX45695.  
XX New nucleic acid molecules and polypeptides useful for diagnosing,  
PT preventing and treating pathological conditions resulting from bacterial  
PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
PT screening.  
XX Disclosure; SEQ ID NO 4871; 301pp; English.  
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX Sequence 550 AA;  
SQ  
Query Match 77.7%; Score 385.5; DB 8; Length 550;  
Best Local Similarity 81.2%; Pred. No. 4.8e-30;  
Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
QY 1 LKEIDESDSEYVKEGLRVPLOSELVDKQAKLLKLEELSDKIDELDAEIAKNLKKVEDF 60  
DB 144 LKEIDESDSEYVKEGLRVPLOSELVDKQAKLLKLEELSDKIDELDAEIAK-LEKNVEDF 202  
QY 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101  
DB 203 KNSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243  
RESULT 7  
ADRS9223  
ID ADR95223 standard; protein; 550 AA.  
XX ADR95223;  
XX 16-DEC-2004 (first entry)  
XX Novel S. pneumoniae protein sequence, SEQ ID 3858.  
XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;  
KW bacterial infection.  
XX Streptococcus pneumoniae.  
OS US6800744-B1.  
XX 05-OCT-2004.  
XX

30-JUN-1998; 98US-00107433.  
XX 02-JUL-1997; 97US-0051553P.  
PR 12-MAY-1998; 98US-0085131P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Doucette-Stamm LA, Bush D;  
XX WPI; 2004-697205/68.  
DR N-ESDB; ADR92620.  
XX New isolated nucleic acid encoding a Streptococcus pneumoniae  
PT polypeptide, useful for diagnosing, preventing and/or treating  
PT pathological conditions resulting from the bacterial infection.  
XX Disclosure; SEQ ID NO 3858; 151pp; English.  
XX The invention relates to an isolated nucleic acid comprising a sequence  
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its  
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,  
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,  
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,  
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or  
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide  
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is  
CC hybridisable under high stringency conditions to the nucleotide sequence.  
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.  
CC Also included are a recombinant expression vector comprising the isolated  
CC nucleic acid cited above operably linked to a transcription regulatory  
CC element, a cell comprising the recombinant expression vector and a probe  
CC comprising at least 20 consecutive nucleotides of the nucleotide  
CC sequences as cited above. The methods and compositions of the present  
CC invention are useful for the diagnosis, prevention and/or treatment of  
CC pathological conditions resulting from bacterial infection by  
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and  
CC otitis media. The present sequence is one of the 2603 disclosed S.  
CC pneumoniae protein sequences. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.  
XX Sequence 550 AA;  
SQ  
Query Match 77.7%; Score 385.5; DB 8; Length 550;  
Best Local Similarity 81.2%; Pred. No. 4.8e-30;  
Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;  
QY 1 LKEIDESDSEYVKEGLRVPLOSELVDKQAKLLKLEELSDKIDELDAEIAKNLKKVEDF 60  
DB 144 LKEIDESDSEYVKEGLRVPLOSELVDKQAKLLKLEELSDKIDELDAEIAK-LEKNVEDF 202  
QY 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101  
DB 203 KNSNGEQAEQYRAAAEEDLAAKQAELEKTEADLKKAVNEPE 243  
RESULT 8  
AAW14584  
ID AAW14584 standard; protein; 194 AA.  
XX AAW14584;  
XX 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX Streptococcus pneumoniae PspA central region.  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX Streptococcus pneumoniae; strain Db16.  
XX



FH Key Location/Qualifiers  
 FT Misc-difference 61 /note= "unidentified amino acid"

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 FT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Db16.

CC Comparison of the N-terminal and central regions (AAW14533-57 and

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence

CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia

CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical

CC region and the immediate 5' tip of the coding sequence are likely to be

CC the critical sequences for predicting PspA cross-reactions and vaccine

CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 194 AA;

Query Match 77.3%; Score 383.5; DB 2; Length 194;

Best Local Similarity 81.0%; Pred. No. 28-30;

Matches 81; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGLRVLQSELDVQAKLLKLELSDKIDELDAETAKNKKDVEDF 60

Db 1 LKEIDSDSDYVKEGLRVLQSELDVQAKLLKLELSDKIDELDAETAKNKKDVEDF 59

QY 61 QNSGGYSALYLEAKKOLVAKAELEKTEADLKKAVHEP 100

Db 60 KXSDGEQAGQYLAABEDLIAKKAELEQTEADLKKAVNEP 99

RESULT 9

ABW02618

ID ABW02618 standard; protein; 194 AA.

XX AC

XX ABW02618;

XX 12-FEB-2004 (first entry)

XX Db16ac pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1..194

FT /note= "Xaa = Unknown amino acid"

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

XX 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain.  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 64; 121pp; English.

XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Db16ac pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention

XX Sequence 194 AA;

Query Match 77.3%; Score 383.5; DB 7; Length 194;

Best Local Similarity 81.0%; Pred. No. 28-30;

Matches 81; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYVKEGLRVLQSELDVQAKLLKLELSDKIDELDAETAKNKKDVEDF 60

Db 1 LKEIDSDSDYVKEGLRVLQSELDVQAKLLKLELSDKIDELDAETAKNKKDVEDF 59

QY 61 QNSGGYSALYLEAKKOLVAKAELEKTEADLKKAVHEP 100

Db 60 KXSDGEQAGQYLAABEDLIAKKAELEQTEADLKKAVNEP 99

RESULT 10

ABW02609

ID ABW02609 standard; protein; 168 AA.

XX AC

XX ABW02609;

XX 12-FEB-2004 (first entry)

XX L81905c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

KW immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1..168

FT /note= "Xaa = Unknown amino acid"

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.  
 PF 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 XX detecting the presence of Streptococcus pneumoniae or its strain,  
 XX comprises at least two different full length isolated gene encoding  
 XX pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 55; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is L81905c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 168 AA;

Query Match 75.9%; Score 376.5; DB 7; Length 168;  
 Best Local Similarity 79.2%; Pred. No. 8.4e-30;  
 Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYVKEGLRVPVLPQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDF 60  
 DB 1 LKEIDSDSDYVKEGLRVPVLPQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDF 59  
 QY 61 QNSGGYSALYLAEEKDLVAKAELEKTEADLKAVHEPE 101  
 DB 60 KNSDGEQAGYLAAREEDLIAKXAEADLKAVDEPE 100

RESULT 11  
 AAW14575  
 ID AAW14575 standard; protein; 167 AA.  
 XX  
 AC AAW14575;  
 XX

DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX

DE Streptococcus pneumoniae PspA central region.

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.

OS Streptococcus pneumoniae; strain L81905.

XX Key Location/Qualifiers

FT Misc-difference 37 /note= "unidentified amino acid"

FT Misc-difference 41 /note= "unidentified amino acid"

FT Misc-difference 83 /note= "unidentified amino acid"

XX

PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 PI  
 XX WPI; 1997-202002/18.  
 DR  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 XX in vaccines for protecting animals against S.pneumoniae infection.  
 PT  
 XX  
 PS Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain AAL81905.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-Oct-2003 to standardise OS field)  
 XX

SQ Sequence 167 AA;

Query Match 73.2%; Score 363; DB 2; Length 167;  
 Best Local Similarity 78.2%; Pred. No. 1.9e-28;  
 Matches 79; Conservative 5; Mismatches 15; Indels 2; Gaps 2;

QY 1 LKEIDSDSDYVKEGLRVPVLPQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDF 60  
 DB 1 LKEIDSDSDYVKEGLRVPVLPQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDF 58

QY 61 QNSGGYSALYLAEEKDLVAKAELEKTEADLKAVHEPE 101

DB 59 KNSDGEQAGYLAAREEDLIAKXAEADLKAVDEPE 99

RESULT 12  
 AAW14568  
 ID AAW14568 standard; protein; 166 AA.  
 XX  
 AC AAW14568;  
 XX

DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX

DE Streptococcus pneumoniae PspA central region.

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.

OS Streptococcus pneumoniae; strain Bg8743.

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 19.1081 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-9  
Perfect score: 496  
Sequence: 1 LKEIDSESDYVKEGLRVP.....KKALEKTEADLKKAHVEPE 101

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	100.0	101	4	US-09-147-875A-9
2	489	98.6	185	4	US-08-529-055-46
3	489	98.6	8991	4	US-08-714-741-32
4	478.5	96.5	102	2	US-08-710-749-8
5	399.5	80.5	100	4	US-09-147-875A-5
6	392.5	79.1	100	4	US-09-147-875A-2
7	390.5	78.7	98	4	US-09-147-875A-1
8	386.5	77.9	183	4	US-08-529-055-50
9	385.5	77.7	550	4	US-09-583-110-4871
10	385.5	77.7	550	4	US-09-107-433-3858
11	384.5	77.3	100	4	US-09-147-875A-3
12	383.5	77.3	194	4	US-08-529-055-64
13	382.5	77.1	100	4	US-09-147-875A-4
14	382	77.0	101	2	US-08-710-749-4
15	376.5	75.9	168	4	US-08-529-055-55
16	375	75.6	101	2	US-08-710-749-1
17	373.5	75.3	100	4	US-09-147-875A-6
18	373	75.2	99	2	US-08-710-749-9
19	372	75.0	101	2	US-08-710-749-3
20	368	74.2	101	2	US-08-710-749-2
21	363	73.2	101	2	US-08-710-749-5
22	354.5	71.5	166	4	US-08-529-055-48
23	346.5	69.9	100	4	US-09-147-875A-8
24	339.5	68.4	100	4	US-09-147-875A-7
25	334	67.3	101	2	US-08-710-749-7
26	329	66.3	99	2	US-08-710-749-10
27	329	66.3	99	4	US-09-147-875A-11

28	329	66.3	204	4	US-08-529-055-51	Sequence 51, Appl
29	327	65.9	101	2	US-08-710-749-6	Sequence 6, Appl
30	322.5	65.0	100	4	US-09-147-875A-12	Sequence 12, Appl
31	313	63.1	99	2	US-08-710-749-11	Sequence 11, Appl
32	313	63.1	198	4	US-08-529-055-61	Sequence 61, Appl
33	313	63.1	619	1	US-08-465-746-2	Sequence 2, Appl
34	313	63.1	619	1	US-08-214-164-2	Sequence 2, Appl
35	313	63.1	619	2	US-08-467-852A-3	Sequence 3, Appl
36	313	63.1	619	2	US-08-246-636-2	Sequence 3, Appl
37	313	63.1	619	2	US-08-247-491A-3	Sequence 3, Appl
38	313	63.1	619	2	US-08-319-795-2	Sequence 2, Appl
39	313	63.1	619	2	US-08-468-985-2	Sequence 2, Appl
40	313	63.1	619	3	US-08-312-949-2	Sequence 2, Appl
41	313	63.1	648	1	US-08-072-070-2	Sequence 2, Appl
42	313	63.1	648	1	US-08-469-434-2	Sequence 2, Appl
43	313	63.1	648	1	US-08-214-222-2	Sequence 2, Appl
44	313	63.1	648	2	US-08-467-852A-2	Sequence 2, Appl
45	313	63.1	648	2	US-08-468-718-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-147-875A-9  
; Sequence 9, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-9

Query Match 100.0%; Score 496; DB 4; Length 101;  
Best Local Similarity 100.0%; Pred. No. 8e-42;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRVPLOQSELDVVKQAKLLLEELSDKIDELDAETAKNLKKDVEDF 60  
Db 1 LKEIDSESDYVKEGLRVPLOQSELDVVKQAKLLLEELSDKIDELDAETAKNLKKDVEDF 60

QY 61 QNSGGYGYSALYLEAAEKDOLVAKKAELEKTEADLKKAHVEPE 101  
Db 61 QNSGGYGYSALYLEAAEKDOLVAKKAELEKTEADLKKAHVEPE 101

RESULT 2  
US-08-529-055-46  
; Sequence 46, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yotter, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York

```

; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529.055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-46

Query Match 98.6%; Score 489; DB 4; Length 185;
Best Local Similarity 99.0%; Pred. No. 8.2e-41;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDESDSYVKEGLRVPLQSELDVYKQAKLLKLELSKIDELDAEIAKNLKKQVEDF 60
Db 1 LKEIDESDSYVKEGLRVPLQSELDVYKQAKLLKLELSKIDELDAEIAKNLKKQVEDF 60

Qy 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
Db 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVNEPE 101

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500813
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996

; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529.055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-46

Query Match 98.6%; Score 489; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 9.6e-39;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDESDSYVKEGLRVPLQSELDVYKQAKLLKLELSKIDELDAEIAKNLKKQVEDF 60
Db 4382 LKEIDESDSYVKEGLRVPLQSELDVYKQAKLLKLELSKIDELDAEIAKNLKKQVEDF 4441

Qy 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
Db 4442 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVNEPE 4482

RESULT 4
US-08-710-749-8
; Sequence 8, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-8
```

```
Query Match      96.5%; Score 478.5; DB 2; Length 102;
Best Local Similarity 98.0%; Pred. No. 4.3e-40;
Matches 100; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLQSEL-DVKAQKLLKLELSKIDELDAEIAKLNKKVDVF 59
    |||||
Db 1 LKEIDSDSDYVKEGLRVPLQSELDDVKAQKLLKLELSKIDELDAEIAKLNKKVDVF 60
    |||||

Qy 60 FONGSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
    |||||
Db 61 FONGSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVNEPE 102
    |||||

RESULT 5
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match      80.5%; Score 399.5; DB 4; Length 100;
Best Local Similarity 83.2%; Pred. No. 2.6e-32;
Matches 84; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLQSEL-DVKAQKLLKLELSKIDELDAEIAKLNKKVDVF 60
    |||||
Db 1 LKEIDSDSDYVKEGERAPLQSELDKQAKLSKLELSKIDELDAEIAK-LEKDVDF 59
    |||||

Qy 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
    |||||
Db 60 KNSDGEQAGYLAARAEEDLIAKAELEKTEADLKKAVHEPE 100
    |||||

RESULT 6
US-09-147-875A-2
; Sequence 2, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-2

Query Match      79.1%; Score 392.5; DB 4; Length 100;
Best Local Similarity 82.2%; Pred. No. 1.3e-31;
Matches 83; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLQSEL-DVKAQKLLKLELSKIDELDAEIAKLNKKVDVF 60
    |||||
Db 1 LKEIDSDSDYVKEGLRVPLQSELDKQAKLSKLELSKIDELDAEIAK-LEKDVDF 59
    |||||

Qy 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
    |||||
Db 60 KNSNGEQAGYRAARAEEDLAAKQAELEKTEADLKKAVHEPE 100
    |||||
```

## RESULT 7

```
US-09-147-875A-1
; Sequence 1, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-1

Query Match      78.7%; Score 390.5; DB 4; Length 98;
Best Local Similarity 86.1%; Pred. No. 2e-31;
Matches 87; Conservative 3; Mismatches 8; Indels 3; Gaps 3;

Qy 1 LKEIDSDSDYVKEGLRVPLQSEL-DVKAQKLLKLELSKIDELDAEIAKLNKKVDVF 60
    |||||
Db 1 LKEIDSDSDYVKEGLRVPLQSELDKQAKLSKLELSKIDELDAEIAK-LEKDVDF 59
    |||||

Qy 61 QNSGGYSALYLEAAEKDVLVAKAELEKTEADLKKAVHEPE 101
    |||||
Db 60 KNS-DGEQAGYLAARAEEDL-AKAELEKTEADLKKAVHEPE 98
    |||||

RESULT 8
US-08-529-055-50
; Sequence 50, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-50

Query Match          77.9%; Score 386.5; DB 4; Length 183;
Best Local Similarity 81.2%; Pred. No. 1e-30;
Matches 82; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAKNIKKQVEDF 60
Db 1 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAK-LEKQVEDF 59

Qy 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 60 QNSDGEQAGQYLAAGGDLIAKKALEKAEADLKKAVDEPE 100

RESULT 9
US-09-583-110-4871
; Sequence 4871, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: PATH00-07A
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4871
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4871

Query Match          77.7%; Score 385.5; DB 4; Length 550;
Best Local Similarity 81.2%; Pred. No. 5.1e-30;
Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAKNIKKQVEDF 60
Db 144 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAK-LEKQVEDF 202

Qy 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 203 KNSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 10
US-09-107-433-3858
; Sequence 3858, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
```

```
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3858:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...550
; SEQUENCE DESCRIPTION: SEQ ID NO: 3858:
US-09-107-433-3858

Query Match          77.7%; Score 385.5; DB 4; Length 550;
Best Local Similarity 81.2%; Pred. No. 5.1e-30;
Matches 82; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAKNIKKQVEDF 60
Db 144 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAK-LEKQVEDF 202

Qy 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 203 KNSNGEQAEQYRAAEEDLAAKQAELEKTEADLKKAVNEPE 243

RESULT 11
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-3

Query Match          77.5%; Score 384.5; DB 4; Length 100;
Best Local Similarity 81.2%; Pred. No. 7.9e-31;
Matches 82; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYVKEGLRVPLOQSLDVQKAKLKLKLELSDKIDELDAEIAKNIKKQVEDF 60
```





Db 1 LKDEIDSESDYVKEGERAPLOQSELDKQAKLSKLELSKIDELDAEIAK-LEKQVED 59  
Qy 60 FQNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAHVEPE 101  
Db 60 FKNSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVEPE 101

RESULT 15  
US-08-529-055-55  
; Sequence 55, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Thereof, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 168 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-55

Query Match 75.9%; Score 376.5; DB 4; Length 168;  
Best Local Similarity 79.2%; Pred. NO. 9.2e-30;  
Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;  
Qy 1 LKDEIDSESDYVKEGLRVLQSELDVQAKLLKLELSKIDELDAEIAKLNKKQVEDF 60  
Db 1 LKDEIDSESDYVKEGFRAPLOQSELDKQAKLSKLELSKIDELDAEIAK-LEKQVEDF 59  
Qy 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAHVEPE 101  
Db 60 KNSDGEQAGQYLAAAEEDLIAKKAELQTEADLKKAHVEPE 100

Search completed: June 18, 2005, 17:07:08  
Job time : 20.1081 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 63.5926 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-9

Perfect score: 496

Sequence: 1 LKEIDSESDYVKEGLRVP.....KKAELEKTEADLKKAHPE 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	100.0	101	US-10-674-755-9	Sequence 9, Appli
2	489	98.6	185	US-10-299-636-61	Sequence 61, Appl
3	399.5	80.5	100	US-10-674-755-5	Sequence 5, Appli
4	392.5	79.1	100	US-10-674-755-2	Sequence 2, Appli
5	390.5	78.7	98	US-10-674-755-1	Sequence 1, Appli
6	386.5	77.9	183	US-10-299-636-65	Sequence 65, Appl
7	384.5	77.5	100	US-10-674-755-3	Sequence 3, Appli
8	383.5	77.3	194	US-10-299-636-79	Sequence 79, Appl
9	382.5	77.1	100	US-10-674-755-4	Sequence 4, Appli
10	376.5	75.9	168	US-10-299-636-70	Sequence 70, Appl
11	373.5	75.3	100	US-10-674-755-6	Sequence 6, Appli

12	354.5	71.5	166	15	US-10-299-636-63	Sequence 63, Appl
13	346.5	69.9	100	15	US-10-674-755-8	Sequence 8, Appli
14	339.5	68.4	100	15	US-10-674-755-7	Sequence 7, Appli
15	329	66.3	99	15	US-10-674-755-11	Sequence 11, Appl
16	329	66.3	204	15	US-10-299-636-66	Sequence 66, Appl
17	322.5	65.0	100	15	US-10-674-755-12	Sequence 12, Appl
18	313	63.1	198	15	US-10-299-636-76	Sequence 76, Appl
19	313	63.1	354	15	US-10-299-636-105	Sequence 105, Appl
20	313	63.1	588	15	US-10-299-636-96	Sequence 96, Appl
21	313	63.1	619	10	US-09-882-774-1	Sequence 1, Appli
22	313	63.1	619	15	US-10-282-122A-73702	Sequence 73702, A
23	313	63.1	619	16	US-10-414-532-72	Sequence 72, Appl
24	311	62.7	170	15	US-10-299-636-75	Sequence 75, Appl
25	311	62.7	181	15	US-10-299-636-57	Sequence 57, Appl
26	311	62.7	643	15	US-10-299-636-95	Sequence 95, Appl
27	311	62.7	670	9	US-09-748-875-63	Sequence 63, Appl
28	311	62.7	670	10	US-09-298-523B-63	Sequence 63, Appl
29	311	62.7	690	9	US-09-748-875-61	Sequence 61, Appl
30	311	62.7	690	10	US-09-298-523B-61	Sequence 61, Appl
31	311	62.7	691	9	US-09-748-875-51	Sequence 1, Appli
32	311	62.7	691	10	US-09-298-523B-1	Sequence 1, Appli
33	311	62.7	701	9	US-09-748-875-62	Sequence 62, Appl
34	311	62.7	701	10	US-09-298-523B-62	Sequence 62, Appl
35	311	62.7	707	9	US-09-748-875-2	Sequence 2, Appli
36	311	62.7	707	10	US-09-298-523B-2	Sequence 2, Appli
37	311	62.7	711	9	US-09-748-875-3	Sequence 3, Appli
38	311	62.7	711	10	US-09-298-523B-3	Sequence 3, Appli
39	311	62.7	739	17	US-10-732-923-3294	Sequence 3294, Ap
40	311	62.7	929	9	US-09-748-875-60	Sequence 60, Appl
41	311	62.7	929	10	US-09-298-523B-60	Sequence 60, Appl
42	311	62.7	929	15	US-10-299-636-94	Sequence 94, Appl
43	308	62.1	99	15	US-10-674-755-13	Sequence 13, Appl
44	308	62.1	188	15	US-10-299-636-74	Sequence 74, Appl
45	304	61.3	99	15	US-10-674-755-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-10-674-755-9  
; Sequence 9, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-9

Query Match 100.0%; Score 496; DB 15; Length 101;  
Best Local Similarity 100.0%; Pred. No. 5.9e-37;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGLRVPLOQSELDVVKQAKLLKLELSDKTIDELDAIKNLKKVDVDF 60  
DB 1 LKEIDSESDYVKEGLRVPLOQSELDVVKQAKLLKLELSDKTIDELDAIKNLKKVDVDF 60  
QY 61 QNSGGVSALYLEAAEKDLVAKAELEKTEADLKKAHPE 101  
DB 61 QNSGGVSALYLEAAEKDLVAKAELEKTEADLKKAHPE 101

RESULT 2

```
US-10-299-636-61
; Sequence 61, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwln
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-61

Query Match      98.6%; Score 489; DB 15; Length 185;
Best Local Similarity 99.0%; Pred. No. 5.1e-36;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVDF 60
Db 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVDF 60

Qy 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101

RESULT 3
US-10-674-755-5
; Sequence 5, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-5

Query Match      80.5%; Score 399.5; DB 15; Length 100;
Best Local Similarity 83.2%; Pred. No. 2.5e-28;
Matches 84; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVDF 60
Db 1 LKEIDESDSEYVKEGERAPLQSELDVQAKLLKLEELSDKIDELDAEIAK-LEKVDVDF 59

Qy 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNSDGEQAGYLAEEEDLIAKAELEKTEADLKKAVHEPE 100

US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```
RESULT 4
US-10-674-755-2
; Sequence 2, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-2

Query Match      79.1%; Score 392.5; DB 15; Length 100;
Best Local Similarity 82.2%; Pred. No. 1.1e-27;
Matches 83; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVDF 60
Db 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAK-LEKNVDF 59

Qy 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNSNGEQAQYRAAEEDLIAKAELEKTEADLKKAVHEPE 100

RESULT 5
US-10-674-755-1
; Sequence 1, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-1

Query Match      78.7%; Score 390.5; DB 15; Length 98;
Best Local Similarity 86.1%; Pred. No. 1.6e-27;
Matches 87; Conservative 3; Mismatches 8; Indels 3; Gaps 3;

Qy 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVDF 60
Db 1 LKEIDESDSEYVKEGLRVPLQSELDVQAKLLKLEELSDKIDELDAEIAK-LEKVDVDF 59

Qy 61 QNSGGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNS-DGEQAQYLAEEEDL-AKAELEKTEADLKKAVHEPE 98

RESULT 6
US-10-299-636-65
; Sequence 65, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```

: GENERAL INFORMATION:
: APPLICANT: Briles, David E
: APPLICANT: McDaniel, Larry S
: APPLICANT: Swiatlo, Edwin
: APPLICANT: Yother, Janet
: APPLICANT: Crain, Marilyn J
: APPLICANT: Hollingshead, Susan
: APPLICANT: Tart, Rebecca
: APPLICANT: Brooks-Walter, Alexis
: TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
: FILE REFERENCE: 57909/361
: CURRENT APPLICATION NUMBER: US/10/299,636
: CURRENT FILING DATE: 2002-11-19
: PRIOR APPLICATION NUMBER: 08/714,741
: PRIOR FILING DATE: 1996-09-16
: PRIOR APPLICATION NUMBER: 08/529,055
: PRIOR FILING DATE: 1995-09-15
: NUMBER OF SEQ ID NOS: 111
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 79
: LENGTH: 194
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (61)
: OTHER INFORMATION: Xaa at position 61 is unknown
US-10-299-636-79

Query Match 77.3%; Score 383.5; DB 15; Length 194;
Best Local Similarity 81.0%; Pred. No. 1.5e-26;
Matches 81; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 1 LKIEDSDSDYVKEGLRVPQSELDVQKAKLKLEELSDKIDELDAEIAKNLKKVDVF 60
DB 1 LKIEDSDSDYVKEGLRVPQSELDVQKAKLKLEELSDKIDELDAEIAKNLKKVDVF 59

QY 61 QNSGGGYSALYLEAAEKDLVAKKAELKTEADLKKAVHEP 100
DB 60 KXSDGQAGYLAAGEDLIAKKAELQTEADLKKAVNEP 99

RESULT 9
US-10-674-755-4
: Sequence 4, Application US/10674755
: Publication No. US20040067237A1
: GENERAL INFORMATION:
: APPLICANT: BECKER et al.
: TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
: FILE REFERENCE: 454312-2471
: CURRENT APPLICATION NUMBER: US/10/674,755
: CURRENT FILING DATE: 2003-09-30
: PRIOR APPLICATION NUMBER: US/09/147,875A
: PRIOR FILING DATE: 1999-05-24
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 100
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-10-674-755-4

Query Match 77.1%; Score 382.5; DB 15; Length 100;
Best Local Similarity 81.2%; Pred. No. 8.5e-27;
Matches 82; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 LKIEDSDSDYVKEGLRVPQSELDVQKAKLKLEELSDKIDELDAEIAKNLKKVDVF 60
DB 1 LKIEDSDSDYVKEGLRVPQSELDVQKAKLKLEELSDKIDELDAEIAKNLKKVDVF 59

QY 61 QNSGGGYSALYLEAAEKDLVAKKAELKTEADLKKAVHEPE 101
DB 60 KNSDGQAGYLAAGEDLIAKKAELKAEADLKKAVDSPE 100

```

```
RESULT 10
US-10-299-636-70
; Sequence 70, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6
Query Match 75.9%; Score 376.5; DB 15; Length 168;
Best Local Similarity 79.2%; Pred. No. 5.4e-26;
Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 60
Db 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 59
Qy 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNSDGEQAGQYLAAAEEDLIAKKAEADLKKAVDEPE 100
RESULT 11
US-10-674-755-6
; Sequence 6, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (38)
; FEATURE:
; OTHER INFORMATION: Xaa at position 38 is unknown
; NAME/KEY: UNSURE
; LOCATION: (42)
; FEATURE:
; OTHER INFORMATION: Xaa at position 42 is unknown
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-70
Query Match 75.9%; Score 376.5; DB 15; Length 168;
Best Local Similarity 79.2%; Pred. No. 5.4e-26;
Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 60
Db 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 59
Qy 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNSDGEQAGQYLAAAEEDLIAKKAEADLKKAVDEPE 100
RESULT 12
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: UNSURE
; LOCATION: (38)
; FEATURE:
; OTHER INFORMATION: Xaa at position 38 is unknown
; NAME/KEY: UNSURE
; LOCATION: (42)
; FEATURE:
; OTHER INFORMATION: Xaa at position 42 is unknown
; NAME/KEY: UNSURE
; LOCATION: (84)
; OTHER INFORMATION: Xaa at position 84 is unknown
US-10-299-636-63
Query Match 71.5%; Score 354.5; DB 15; Length 166;
Best Local Similarity 76.2%; Pred. No. 5e-24;
Matches 77; Conservative 7; Mismatches 16; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 60
Db 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 59
Qy 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 PMSDGEQAGQYLVAEEKDLDAKEAEELGNTGADLKKAVDEPE 100
RESULT 13
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
```

```
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-10-674-755-6
Query Match 75.3%; Score 373.5; DB 15; Length 100;
Best Local Similarity 79.2%; Pred. No. 5.4e-26;
Matches 80; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 60
Db 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 59
Qy 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 KNSDGEQAGQYLAAAEEDLIAKKAEADLKKAVDEPE 100
```

```
RESULT 12
US-10-299-636-63
; Sequence 63, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-63
Query Match 71.5%; Score 354.5; DB 15; Length 166;
Best Local Similarity 76.2%; Pred. No. 5e-24;
Matches 77; Conservative 7; Mismatches 16; Indels 1; Gaps 1;
Qy 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 60
Db 1 LKEIDESDSEYVKEGLRVPQSELVDVQAKLLKLEELSDKIDELDAEIAKNLKKVDVF 59
Qy 61 QNSGGYSALYLEAAEKDLVAKAELEKTEADLKKAVHEPE 101
Db 60 PMSDGEQAGQYLVAEEKDLDAKEAEELGNTGADLKKAVDEPE 100
```

```
RESULT 13
US-10-674-755-8
; Sequence 8, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
```

;; PRIOR FILING DATE: 1999-05-24  
;; NUMBER OF SEQ ID NOS: 28  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 8  
;; LENGTH: 100  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-8

Query Match 69.9%; Score 346.5; DB 15; Length 100;  
Best Local Similarity 75.2%; Pred. No. 1.4e-23;  
Matches 76; Conservative 6; Mismatches 18; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 60  
DB 1 LKGIIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 60  
QY 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101  
DB 60 KLTDAEQTEQYLAAAEKDLADKKALEKTEADLKKAVHEPE 100

RESULT 14  
US-10-674-755-7  
;; Sequence 7, Application US/10674755  
;; Publication No. US20040067237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BECKER et al.  
;; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
;; FILE REFERENCE: 454312-2471  
;; CURRENT APPLICATION NUMBER: US/10/674,755  
;; CURRENT FILING DATE: 2003-09-30  
;; PRIOR APPLICATION NUMBER: US/09/147,875A  
;; PRIOR FILING DATE: 1999-05-24  
;; NUMBER OF SEQ ID NOS: 28  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 7  
;; LENGTH: 100  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
;; FEATURE:  
;; NAME/KEY: UNSURE  
;; LOCATION: (1)..(100)  
;; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-7

Query Match 68.4%; Score 339.5; DB 15; Length 100;  
Best Local Similarity 74.3%; Pred. No. 6e-23;  
Matches 75; Conservative 6; Mismatches 19; Indels 1; Gaps 1;  
QY 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 60  
DB 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 59  
QY 61 QNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101  
DB 60 PNSDGEQAGQYLVAEEKDLDAKEAELGNTGADLKKAVDEPE 100

RESULT 15  
US-10-674-755-11  
;; Sequence 11, Application US/10674755  
;; Publication No. US20040067237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BECKER et al.  
;; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
;; FILE REFERENCE: 454312-2471  
;; CURRENT APPLICATION NUMBER: US/10/674,755  
;; CURRENT FILING DATE: 2003-09-30  
;; PRIOR APPLICATION NUMBER: US/09/147,875A  
;; PRIOR FILING DATE: 1999-05-24  
;; NUMBER OF SEQ ID NOS: 28  
;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 11  
;; LENGTH: 99  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-11  
Query Match 66.3%; Score 329; DB 15; Length 99;  
Best Local Similarity 68.8%; Pred. No. 5.2e-22;  
Matches 75; Conservative 5; Mismatches 11; Indels 18; Gaps 3;  
QY 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 54  
DB 1 LKEIDSDSDYVKEGLRVPLOSELVQKAKLLKLEELSDKIDELDAETAKLKKDVEDF 60  
QY 55 --KDVEDFQNSGGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101  
DB 61 ENNVEDY-----FKEGLEKTIAAKKALEKTEADLKKAVNEPE 99  
Search completed: June 18, 2005, 18:00:24  
Job time : 64.5926 secs

**This Page Blank (uspto)**





```
Infect. Immun. 59, 1285-1289, 1991
A>Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A;Reference number: A60282; MUID:91169598; PMID:2004810
A;Accession: A60282
A;Molecule type: protein
A;Residues: 32-76 <TAL>
A;Experimental source: strain JV2008
C;Genetics:
A;Gene: pspA
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-619/Product: surface protein pspA #status predicted <MAT>
F:411-430/Domain: cpl repeat homology <CP01>
F:431-450/Domain: cpl repeat homology <CP02>
F:451-470/Domain: cpl repeat homology <CP03>
F:471-490/Domain: cpl repeat homology <CP04>
F:491-510/Domain: cpl repeat homology <CP05>
F:511-530/Domain: cpl repeat homology <CP06>
F:531-550/Domain: cpl repeat homology <CP07>
F:551-570/Domain: cpl repeat homology <CP08>
F:571-591/Domain: cpl repeat homology <CP09>
F:592-611/Domain: cpl repeat homology <CP10>

Query Match 63.1%; Score 313; DB 2; Length 619;
Best Local Similarity 65.1%; Pred. No. 7; 1e-16;
Matches 71; Conservative 8; Mismatches 12; Indels 18; Gaps 3;

QY 1 LKEIDSESDYVKEGLRVPLQSELDVVKQAKLLKLEELSDKIDELDAIAK---NLK--- 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 LKEIDSESDYVKEGLRVPLQSELDVVKQAKLLKLEELSDKIDELDAIAKLELDQKAAE 282

QY 55 --KVEDFQNSGGYSALYLEAAEKDLVAKAELEKTEADLKAVHEE 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 ENNVEDY-----FKGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 3
F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95013
R;Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <KUR>
A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:g14971584; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0117

Query Match 22.3%; Score 110.5; DB 2; Length 744;
Best Local Similarity 31.8%; Pred. No. 0.74;
Matches 42; Conservative 15; Mismatches 36; Indels 39; Gaps 6;

QY 2 KEIDE-----SDSEDYVKEGLRVPLQSELDVVKQAKLLK-----LEELSDKI-----D 43
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 KEISNLEILLGGADPEDDT-----AALQNKLAAKKAELAKKQTELEKLLDSDPGEKTD 368

QY 44 ELDAETAK-NLKQVDVFQNS-----GGYSALYLEAAEKDLVAKKAELE 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 369 ELQKGAEEELQKKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLATKKAELE 428

QY 88 KTEADLKAVHEE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 429 KTKELDAALNE 440
```

## RESULT 4

```
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34418
R;Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid F12F3.
A;Reference number: Z21521
A;Accession: T34418
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3488 <FUL>
A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
A;Experimental source: strain Bristol N2; clone F12F3
C;Genetics:
A;Gene: CESP:F12F3.3
A;Map position: 5
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 21.0%; Score 104; DB 2; Length 3488;
Best Local Similarity 35.4%; Pred. No. 12;
Matches 35; Conservative 16; Mismatches 36; Indels 12; Gaps 4;

QY 1 LKEIDSESDYVKE-GLRVPLQSELDVVKQKL---LKLEELSDKIDELDAIAKLNKKDV 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1041 IKTVSEDDAARKEKELNDKLESEIATPKASADKLEEQAKAEVAAKKQKEKD 1100

QY 58 EDFQNSGGYSALYLEAAEKDLVAKAELEKTEADLKKA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1101 EQLK-----LQTEAAKKAAAEKLELEK-QAQIKKA 1130

RESULT 5
C70445
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: C70445
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: C70445
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1006 <AQF>
A;Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: clpB
C;Superfamily: endopeptidase Clp ATP-binding chain
C;Keywords: hydrolase

Query Match 20.1%; Score 99.5; DB 2; Length 1006;
Best Local Similarity 33.9%; Pred. No. 6.6;
Matches 37; Conservative 19; Mismatches 28; Indels 25; Gaps 6;

QY 1 LKEIDSEDS-----DYVKEGLRVPLQSELDVVKQAKLLK-LEELSDKIDELDAIAKLNK 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 552 IKALEEQIIEANLKGDIYE-----AQLKIEKAKLEKQELLGKRVGGVKAIAE-LK 603

QY 55 KDVEDFQNSGGYSALYLEAAEKDLVAKKAELE-----KTEADLKAVHEE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 604 KKHLEDEK-----IKAAEKGDYKEAEELKIEKAKLEKELKLEQE 645

RESULT 6
I49464
alpha cardiac myosin heavy chain - mouse
```

C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I49464; I49463; I49462; I49461; I49604  
R:Quinn-Laquer, B.K.; Kennedy, J.E.; Wei, S.J.; Beisel, K.W.  
Genomics 13, 176-188, 1992  
A:Title: Characterization of the allelic differences in the mouse cardiac alpha-myosin h  
A:Reference number: A38207; MUID:92250040; PMID:11577481  
A:Accession: I49464  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1938 <RES>  
A:Cross-references: UNIPROT:Q02566; GB:M76601; NID:g191623; PIDN:AAA37162.1; PID:g191624  
A:Accession: I49463  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-193 'D', 195-837, 'S', 839-955, 'N', 957-1938 <RE2>  
A:Cross-references: GB:M76600; NID:g191621; PIDN:AAA37161.1; PID:g191622  
A:Accession: I49462  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1938 <RES>  
A:Cross-references: GB:M76599; NID:g191619; PIDN:AAA37160.1; PID:g191620  
A:Accession: I49461  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-544, 'A', 546-1938 <RE4>  
A:Cross-references: GB:M76598; NID:g191617; PIDN:AAA37159.1; PID:g191618  
R:Gulick, J.; Subramaniam, A.; Neumann, J.; Robbins, J.  
J. Biol. Chem. 266, 9180-9185, 1991  
A:Title: Isolation and characterization of the mouse cardiac myosin heavy chain genes.  
A:Reference number: I49604; MUID:91225025; PMID:2026617  
A:Accession: I49604  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-67 <RES>  
A:Cross-references: GB:M62404; NID:g192609; PIDN:AAA37424.1; PID:g192610  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: ATP; cardiac muscle; heart; nucleotide binding; P-loop  
P:88-768/Domain: myosin motor domain homology <MMOT>  
P:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match 19.6%; Score 97; DB 2; Length 1938;  
Best Local Similarity 26.6%; Pred.No.20;  
Matches 29; Conservative 31; Mismatches 33; Indels 16; Gaps 5;

Oy 2 KEIDSDSDYV--KEGLRVPLQSELDVKKAKLLKEELSD-----KIDELDARIKN 52  
Db 1085 KEFDISQNSKIDEDQALALQKQKENGARLEEELEAEARTAKVKELRSLRSRE 1144  
Oy 53 LKKDVEDFQNSGGYSALYLEAAEKDIAVKKAKAEKTEADLKKAV--HE 99  
Db 1145 LEEISERLEEAGGA-TSQVIEMNKK-----REAEFQKWRDLDEATLQHE 1188

RESULT 7  
S06005  
N:Alternate names: alpha-myosin heavy chain  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S06005; S07535; A20971; A02988; I53305  
R:McNally, E.M.; Gianola, K.M.; Leinwand, L.A.  
Nucleic Acids Res. 17, 7527-7528, 1989  
A:Title: Complete nucleotide sequence of full length cDNA for rat alpha cardiac myosin h  
A:Reference number: S06005; MUID:90016822; PMID:2798111  
A:Accession: S06005  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-1938 <MCN>  
A:Cross-references: UNIPROT:P02563; EMBL:X15938; NID:g56654; PIDN:CAA34064.1; PID:g56659  
R:McNally, E.M.; Kratt, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.  
J. Mol. Biol. 210, 665-671, 1989

A>Title: Full-length rat alpha and beta cardiac myosin heavy chain sequences. Compared  
A/Reference number: S07535; MUID:90133919; PMID:2614840  
A/Accession: S07535  
A>Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-1938 <MC2>  
R:Mandavi, V.; Chambers, A.P.; Nadal-Ginard, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 2626-2630, 1984  
A>Title: Cardiac alpha- and beta-myosin heavy chain genes are organized in tandem.  
A/Reference number: A20971; MUID:84194059; PMID:6585819  
A/Accession: A20971  
A/Molecule type: protein  
A/Residues: 1-12, 'AP', 'I4-45', 'A', '47-50', 'AP', '53-81', 'E', '83-86', 'Q', '88-109, 'I11-133', 'H', '135-1  
R:Mandavi, V.; Periasamy, M.; Nadal-Ginard, B.  
Nature 297, 659-664, 1982  
A>Title: Molecular characterization of two myosin heavy chain genes expressed in the ad  
A/Reference number: A02988; MUID:82220036; PMID:7045682  
A/Accession: A02988  
A/Molecule type: mRNA  
A/Residues: 1512-1574  
A/Note: there are 10 or more myosin heavy chain genes in the rat, at least two of which  
R:Mandavi, V.; Lompre, A.M.; Chambers, A.P.; Nadal-Ginard, B.  
Eur. Heart J. 5, 181-191, 1984  
A>Title: Cardiac myosin heavy chain isozytic transitions during development and under p  
A/Reference number: I53305; MUID:85179510; PMID:6241892  
A/Accession: I53305  
A>Status: translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1872-1933, 'I', '1935-1938 <RES>  
A/Cross-references: GB:M32697; NID:g205596; PIDN:AAA41658.1; PID:g205597  
C/Superfamily: myosin heavy chain; myosin motor domain homology  
C/Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methylat  
F:87-767/Domain: myosin motor domain homology <MMOT>  
F:177-184/Region: nucleotide-binding motif A (P-loop)  
F:548-585/Region: actin binding #status predicted  
F:656-678/Region: actin binding #status predicted  
F:840-1938/Domain: coiled coil #status predicted <COI>  
F:840-1280/Region: S2  
F:1281-1938/Region: light meromyosin  
F:128/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:183/Binding site: ATP (Lys) #status predicted  
F:696,706/Active site: Cys #status predicted

Query Match 19.4%; Score 96; DB 1; Length 1938;  
Best Local Similarity 26.6%; Pred. No. 24;  
Matches 29; Conservative 30; Mismatches 34; Indels 16; Gaps 5;

Oy 2 KEIDESDSEYV--KEGRVPLOSELVDVKQAKLLLEELSD-----KIDELDAETAKN 52  
. ||| :  
Db 1084 KEFDISQNSKIETDEQALQLQKKLKENQARIEELEEEAEARTAKVEKLRSDLTRE 1143  
||| :

Oy 53 LKKOVEDFQNSGGYSALYLEAAEKDLVAKKAELKTEADLKAV--HE 99  
||| : ||  
Db 1144 LEETSERLEEAGGA-TSQVIENMK---REAEFQKMRRDLLEATLQHE 1187  
||| : ||

RESULT 8  
148175  
myosin heavy chain alpha, cardiac muscle [similarity] - golden hamster  
C/Species: Mesocricetus auratus (golden hamster)  
C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C/Accession: I48175; A23938  
R/Wang, R.; Sole, M.J.; Cukerman, E.; Liew, C.C.  
J. Mol. Cell. Cardiol. 26, 1155-1165, 1994  
A>Title: Characterization and nucleotide sequence of the cardiac alpha-myosin heavy cha  
A/Reference number: I48153; MUID:95115033; PMID:7815459  
A/Accession: I48175  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-1939 <RES>  
A/Cross-references: UNIPROT:P13539; GB:I15351; NID:g402373; PIDN:AAB59701.1; PID:g40237  
R/Liew, C.C.; Jandreski, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3175-3179, 1986





C:Species: Plasmodium vivax  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 28-Apr-1995  
C:Accession: A42771  
R:Galinski, M.R.; Medina, C.C.; Ingravall, P.; Barnwell, J.W.  
Cell 69, 1213-1226, 1992  
A:Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.  
A:Reference number: A42771; MUID:92315338; PMID:1617731  
A:Accession: A42771  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2829 <GAL>  
A:Experimental source: Belem strain, merozoites  
A>Note: sequence extracted from NCBI backbone (NCBIN:108114, NCBIPI:108115)

Query Match 18.5%; Score 92; DB 2; Length 2829;  
Best Local Similarity 30.8%; Pred. No. 71;  
Matches 33; Conservative 22; Mismatches 32; Indels 20; Gaps 5;

QY 1 LKEIDSESDYVK--EGLRVPLOSELDVQAKLLKLELSKDIDEL---DAETAKNLKK 55  
DB 1418 VKLDYSIDIDKKVKEIGK---REILKMKESALTFWEE-SEKFKQMCSSHENAKGKK 1473

QY 56 DVEDFQNSGGGYSA-----LYLEAAEKDLVAKKAELEKTEA 91  
DB 1474 KIEYLKNGDGGKANITDSQMEEVGVYVSKAEHAFHTVEAQVDKTKA 1520

RESULT 15  
C64239  
ATP-dependent proteinase clpB (regulator component) homolog - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: C64239  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346; PMID:7569993  
A:Accession: C64239  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-714 <TIGR>  
A:Cross-references: UNIPROT:P47597; GB:U39719; GB:L43967; NID:gl046055; PID:gl046062; TI  
A:Experimental source: strain G-37  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: endopeptidase Clp ATP-binding chain  
C:Keywords: nucleotide binding; P-loop  
F:60-67/Region: nucleotide-binding motif A (P-loop)  
F:466-473/Region: nucleotide-binding motif A (P-loop)

Query Match 18.2%; Score 90.5; DB 2; Length 714;  
Best Local Similarity 31.8%; Pred. No. 21;  
Matches 35; Conservative 19; Mismatches 33; Indels 23; Gaps 6;

QY 1 LKEIDSESDYVKGLRVPLOSELD-VQAKLLKLELSKDIDELDAETAKNLKQVED 59  
DB 285 LKQDKENDNKQSKKEYLE-KLKQLDALKQRDSLINEW--KKEKADFENINKLKEIEE 341

QY 60 FQ-----NSGGGYSA---LY-----LEAAEKDLVAKKAELEKTE 90  
DB 342 FQTKLETYQSEGNYSASKILYSDIPRLKKELESQAQQKYATSKHDLFKTE 391

Search completed: June 18, 2005, 17:03:56  
Job time : 14.2431 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	464.5	93.6	393	2	Q9LAZ3	streptococ
2	443.5	89.4	255	2	Q9L581	streptococ
3	443.5	89.4	255	2	Q9L586	streptococ
4	424.5	85.6	246	2	Q9L578	streptococ
5	402.5	81.1	340	2	Q8KQK5	streptococ
6	395.5	79.7	406	2	Q9LAZ0	streptococ
7	392.5	79.1	394	2	Q9LAY6	streptococ
8	392.5	79.1	395	2	Q9LAZ1	streptococ
9	386.5	77.9	225	2	Q9L591	streptococ
10	385.5	77.7	222	2	Q9L577	streptococ
11	385.5	77.7	262	2	Q9L576	streptococ
12	385.5	77.7	415	2	Q9LAY7	streptococ
13	379.5	76.5	194	2	Q9L5B5	streptococ
14	379.5	76.5	218	2	Q6UEB2	streptococ
15	379.5	76.5	233	2	Q9L568	streptococ
16	379.5	76.5	236	2	Q9L569	streptococ
17	379.5	76.5	243	2	Q9L564	streptococ
18	379.5	76.5	243	2	Q9L567	streptococ
19	379.5	76.5	244	2	Q9L565	streptococ
20	379.5	76.5	247	2	Q9L566	streptococ
21	379.5	76.5	249	2	Q9L570	streptococ
22	379.5	76.5	254	2	Q9L563	streptococ
23	379.5	76.5	401	2	Q9LAZ2	streptococ
24	378.5	76.3	416	2	Q9LAY8	streptococ
25	359.5	72.5	207	2	Q8GNS9	streptococ
26	347.5	70.1	237	2	Q9L592	streptococ
27	347.5	70.1	395	2	Q9LAY9	streptococ
28	329	66.3	417	2	Q9LAY3	streptococ
29	313	63.1	437	2	Q9LAY4	streptococ
30	313	63.1	619	2	Q54972	streptococ
31	313	63.1	619	2	Q8DR10	streptococ

```

RP SEQUENCE FROM N.A.
RC STRAIN=SP221;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP221;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255546; AAF68099.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR TrEMBL; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 255
FT NON_TER 255
SQ SEQUENCE 255 AA; 28524 MW; 0D855A5B2DAS9A27 CRC64;

Query Match 89.4%; Score 443.5; DB 2; Length 255;
Best Local Similarity 93.1%; Pred. No. 2.4e-23;
Matches 94; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAKLNKKVDVF 60
Db 65 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAK-LKQVDVF 123

Qy 61 QNSGGGSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 124 KNSDGEISALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 164

RESULT 3
Q9LSB6 PRELIMINARY; PRT; 255 AA.
AC Q9LSB6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP200;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253406; AAF67354.1; -.
DR InterPro; IPR009082; His_kin_homodim.
DR TrEMBL; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 255
FT NON_TER 255
SQ SEQUENCE 255 AA; 28496 MW; 0D855A4210135877 CRC64;

Query Match 89.4%; Score 443.5; DB 2; Length 255;
Best Local Similarity 93.1%; Pred. No. 2.4e-23;
Matches 94; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 1 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAKLNKKVDVF 60
Db 65 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAK-LKQVDVF 123

Qy 61 QNSGGGSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 124 KNSDGEISALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 164

RESULT 4
Q9LS78 PRELIMINARY; PRT; 246 AA.
AC Q9LS78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=237;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=237;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255549; AAF68102.1; -.
DR HSSP; P00192; 1Q03.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 246
FT NON_TER 246
SQ SEQUENCE 246 AA; 27321 MW; E4459259C5518656 CRC64;

Query Match 85.6%; Score 424.5; DB 2; Length 246;
Best Local Similarity 89.1%; Pred. No. 4.8e-24;
Matches 90; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAKLNKKVDVF 60
Db 55 LKEIDSESDYVKEGLRVPLQSELDVQAKLLKLELSDKIDELDAEIAK-LKQVDVF 113

Qy 61 QNSGGGSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
Db 114 KNSDGEISALYLEAAEKDLVAKKAELEKTEADLKKAVNEPE 154

RESULT 5
Q8KQK5 PRELIMINARY; PRT; 340 AA.
AC Q8KQK5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=St 435/96;
RX MEDLINE=22170754; PubMed=12183557;

```



RX	DOI=10.1128/TAI.70.9.5086-5090.2002;	
RA	Myaaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,	
RA	Das W.O., Leite L.C.C.;	
RT	"Analysis of serum cross-reactivity and cross-protection elicited by	
RT	immunization with DNA vaccines against Streptococcus pneumoniae	
RT	expressing PcpA fragments from different clades.";	
RL	Infect. Immun. 70:5086-5090(2002).	
DR	EMBL; AY082387; AAU92492.1; -	
DR	InterPro; IPR009082; His kin homodim.	
DR	InterPro; IPR005033; Tropomyosin.	
DR	PRINTS; PR00194; TROPOMYOSIN.	
FT	NON TER 1	
FT	NON TER 340 340	
SQ	SEQUENCE 340 AA; 38023 MW; EE07ECF00B1FBD57 CRC64;	
	Query Match 81.1%; Score 402.5; DB 2; Length 340;	
	Best Local Similarity 85.1%; Pred. No. 2.2e-20;	
	Matches 86; Conservative 5; Mismatches 9; Indels 1; Gaps 1;	
Qy	1 LKEIDESDSEDYVVEGRLVPLQSELDVKQAKLLKLELSDKIDELDAEIAKLNKKQVEDF 60	
Db	197 LKEIDESDSEDYVVEGRLVPLQSELDVKQAKLLKLELSDKIDELDAEIAK-LKKQVEDF 255	
Qy	61 QNSGGGYSAlyLEAAEKDVLVAKKAELKTEADLKKAVHEPE 101	
Db	256 KNSDGEQAGYGLAAAEEDLVAKKAELKTEADLKKAVNEPE 296	
RESULT 6		
Q9LAZ0		
ID	Q9LAZ0 PRELIMINARY; PRT; 406 AA.	
AC	Q9LAZ0;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	PcpA (Fragment).	
GN	Name=pcpA;	
OS	Streptococcus pneumoniae;	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1313;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=DBL6A;	
RX	MEDLINE=20448953; PubMed=10992499;	
RX	DOI=10.1128/TAI.68.10.5889-5900.2000;	
RA	Hollingshead S.K., Becker R., Briles D.E.;	
RT	"Diversity of PcpA: mosaic genes and evidence for past recombination	
RT	in Streptococcus pneumoniae.";	
RL	Infect. Immun. 68:5889-5900(2000).	
DR	EMBL; AF071805; AAP27701.1; -	
DR	InterPro; IPR009082; His kin homodim.	
DR	InterPro; IPR005033; Tropomyosin.	
DR	PRINTS; PR00194; TROPOMYOSIN.	
FT	NON TER 406 406	
SQ	SEQUENCE 406 AA; 591E20CA4D06F052 CRC64;	
	Query Match 79.7%; Score 395.5; DB 2; Length 406;	
	Best Local Similarity 82.2%; Pred. No. 8e-20;	
	Matches 83; Conservative 7; Mismatches 10; Indels 1; Gaps 1;	
Qy	1 LKEIDESDSEDYVVEGRLVPLQSELDVKQAKLLKLELSDKIDELDAEIAKLNKKQVEDF 60	
Db	213 LKEIDESDSEDYVVEGRLVPLQSELDVKQAKLLKLELSDKIDELDAEIAK-LKKQVEDF 271	
Qy	61 QNSGGGYSAlyLEAAEKDVLVAKKAELKTEADLKKAVHEPE 101	
Db	272 KNSDGEQAGYGLAAAEEDLVAKKAELKTEADLKKAVNEPE 312	
RESULT 7		
Q9LAY6		
ID	Q9LAY6 PRELIMINARY; PRT; 394 AA.	



```

RESULT 12
Q9LAY7 PRELIMINARY; PRT; 415 AA.
AC Q9LAY7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PsA (fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG6692;
RX DOI=10.1128/JAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of SpA; mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071808; AAF27704.1; -.
FT NON_TER 415
SQ SEQUENCE 415 AA; 45593 MW; 41375ACBFA10FA46 CRC64;

Query Match 77.7%; Score 385.5; DB 2; Length 415;
Best Local Similarity 81.2%; Pred No. 4e-19; Indels 1; Gaps 1;
Matches 82; Conservative 7; Mismatches 11;

Qy 1 LKEIDESDSYVVEKGLRVLPQSLELDVKQAQLKLLELSDKIDELDAEIAKNLKVDVF 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 229 LKEIDESDSYVVEKGLRAPQSLELDKAQAKSLKLELSKDIDELDAEIAK-LKQNVDF 287
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 61 QNSGGYSALYLEAAEKDLVAKKAELEKTEADLKKAVHEPE 101
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 288 KNSNGEAQEQYRAAAEBDLAAKQAELEKTEADLKKAVNEPE 328
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 13
Q9LSB5 PRELIMINARY; PRT; 194 AA.
ID Q9LSB5 AC Q9LSB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PsA (fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
disseminated clones.";
J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP196;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF253407; AAF67355.1; -.
InterPro; IPR009082; His_kin_homodim.
FT NON_TER 194
SQ SEQUENCE 194 AA; 21116 MW; E68189FCA2B244F8 CRC64;
```

RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=39;

RA Beall B.W.;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF255902; AAF70092.1; -.

DR InterPro; IPR009082; His\_kin\_homodim.

FT NON\_TER 1

FT NON\_TER 233

SQ SEQUENCE 233 AA; 24514 MW; D5C494019C45BFE2 CRC64;

Query Match 76.5%; Score 379.5; DB 2; Length 233;  
Best Local Similarity 80.2%; Pred. No. 6.1e-19;  
Matches 81; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYVKEGLRVPLOSELDVVKQAKLLKLELSDKIDELDAEIAKNLKKVDF 60

Db 28 LKEIDSESDYVKEGLRVPLOSELDVVKQAKLLKLELSDKIDELDAEIAKNLKKVDF 60

Qy 61 QNSGGYSALYLEAEKDLVAKAELEKTEADLKKVHEPE 101

Db 87 KNSDGEQAEQYLVAAKOLDAKKAELTEADLKKVDEPE 127

Search completed: June 18, 2005, 17:01:36  
Job time : 61.5706 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:31:50 ; Search time 73.0731 Seconds  
(without alignments)  
529.279 Million cell updates/sec

Title: US-10-674-755-10

Perfect score: 494

Sequence: 1 LKIDSDSEYANEGFRAP.....KXAELEKAEADLKKAVDEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	459.5	93.0	170	7	ABW02614
2	459.5	93.0	181	7	ABW02596
3	459.5	93.0	865	6	ABU08489
4	459.5	93.0	929	2	AAW14593
5	459.5	93.0	929	2	AAW14593
6	459.5	93.0	8991	6	ABU08487
7	456.5	92.4	188	7	AAW14580
8	456.5	92.4	188	7	ABW02613
9	447.5	90.6	588	6	ABU08491
10	447.5	90.6	589	2	AAW14392
11	445.5	90.2	204	2	AAW14578
12	445.5	90.2	204	2	ABW02612
13	445	90.1	180	2	AAW14562
14	442	89.5	187	2	AAW14579
15	439.5	89.0	1231	6	ABU08490
16	432.5	87.6	206	2	AAW14574
17	432.5	87.6	206	2	ABW02608
18	423.5	85.7	198	2	AAW14581
19	420.5	85.1	198	7	ABW02615
20	420.5	85.1	315	2	AAW04375
21	420.5	85.1	619	2	AAW63437
22	420.5	85.1	619	2	AAW87598
23	420.5	85.1	619	2	AAW86911
24	420.5	85.1	619	2	AAW14838
25	420.5	85.1	619	5	AAE18782

26	420.5	85.1	619	6	ABU45778
27	420.5	85.1	619	8	ADO52126
28	420.5	85.1	648	2	AAW70336
29	420.5	85.1	648	2	AAW62274
30	420.5	85.1	648	2	AAW41837
31	420.5	85.1	648	2	AAW87879
32	420.5	85.1	653	2	AAW92456
33	420.5	85.1	684	2	AAW73912
34	412.5	83.5	204	2	AAW14571
35	412.5	83.5	204	7	ABW02605
36	400.5	81.1	653	2	AAW27150
37	398	80.6	289	2	AAW62276
38	398	80.6	289	2	AAW41840
39	398	80.6	289	2	AAW87910
40	398	80.6	289	2	AAW92458
41	393.5	79.7	195	2	AAW14591
42	393.5	79.7	195	7	ABW02625
43	376.5	76.2	623	6	ABU08494
44	350	70.9	605	6	ABU08493
45	346.5	70.1	190	2	AAW14569

## ALIGNMENTS

### RESULT 1

ABW02614

ID ABW02614 standard; protein; 170 AA.

AC ABW02614;

DT 12-FEB-2004 (first entry)

DE Rct135c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

KW immunological; gene therapy; immunostimulant.

OS Unidentified.

PN US6592876-B1.

PD 15-JUL-2003.

PF 15-SEP-1995; 95US-00529055.

PR 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

(UABR-) UAB RES FOUND.

Briles DE, McDaniel LS, Swiatlo B, Yother J, Brooks-Walter A;

WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for

detecting the presence of Streptococcus pneumoniae or its strain,

comprises at least two different full length isolated gene encoding

pneumococcal surface protein A.

Example 6; SEQ ID NO 60; 121pp; English.

The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Rct135c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 170 AA;

Query Match 93.0%; Score 459.5; DB 7; Length 170;  
Best Local Similarity 96.0%; Pred. No. 1.2e-35;  
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESSEYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDA 60  
Db 1 LKEIDESSEYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLE-VQLKDA 59

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVIDEPE 100  
Db 60 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 2  
ABW02596  
ID ABW02596 standard; protein; 181 AA.  
XX  
AC ABW02596;  
XX  
AC ABW02596;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE 0922134c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
FN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
PI WPI; 2003-862841/80.  
XX  
DR Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
PS Example 6; SEQ ID NO 42; 121bp; English.  
XX

The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspA) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is 0922134c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
SQ Sequence 181 AA;

Query Match 93.0%; Score 459.5; DB 7; Length 181;  
Best Local Similarity 96.0%; Pred. No. 1.3e-35;  
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESSEYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDA 60  
Db 1 LKEIDESSEYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLE-VQLKDA 59

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVIDEPE 100  
Db 60 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 3  
ABU08489  
ID ABU08489 standard; protein; 865 AA.  
XX  
AC ABU08489;  
XX  
DT 24-JUN-2003 (first entry)  
XX  
DE S. pneumoniae pneumococcal surface protein C (PspC) protein.  
XX  
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
KW antibacterial.  
XX  
OS Streptococcus pneumoniae.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..37  
FT /label= Signal\_peptide  
FT Protein 38..865  
FT /label= Mature\_PspC\_protein  
XX  
FN US6500613-B1.  
XX  
PD 31-DEC-2002.  
XX  
PF 16-SEP-1996; 96US-00714741.  
XX  
PR 15-SEP-1995; 95US-00529055.  
XX  
PA (UYAL-) UNIV ALABAMA.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
DR WPI; 2003-361534/34.  
DR N-PSDB; ABX95377.  
XX  
PT Isolated PspC amino acid sequence used as polymerase chain reaction or  
PT hybridization probe, comprises pneumococcal surface protein having alpha-  
PT helical, proline rich and repeat regions.  
XX  
PS Claim 3; Fig 21; 186pp; English.  
XX

The present invention relates to the isolation of Streptococcus  
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
CC like protein having alpha-helical, proline rich and repeat regions. The  
CC PspC and PspA proteins may be used in a vaccine to protect against  
CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
CC PspA may be used for the expression of the proteins, and as PCR primers  
CC or hybridisation probes. The present sequence represents S. pneumoniae  
CC PspC protein  
XX  
SQ Sequence 865 AA;

Query Match 93.0%; Score 459.5; DB 6; Length 865;  
Best Local Similarity 96.0%; Pred. No. 8e-35;  
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKGFRAPLOSQKIDAKKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
 |||||  
 Db 466 LKEIDSDSDYAKGFRAPLOSQKIDAKKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 524  
 |||||  
 QY 61 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAADLKKAVDEPE 100  
 |||||  
 Db 525 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAADLKKAVDEPE 564  
 |||||

## RESULT 4

AAW14593  
 ID AAW14593 standard; protein; 929 AA.  
 XX  
 AC AAW14593;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 27-OCT-1997 (first entry)  
 XX  
 XX Streptococcus pneumoniae PspC.  
 XX  
 XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 XX Streptococcus pneumoniae; strain EF6796.  
 XX

Key Location/Qualifiers  
 FT Peptide 1..37  
 FT /label= Sig\_peptide  
 FT Protein 38..929  
 FT /label= Mat\_protein  
 FT Domain 38..637  
 FT /label= Alpha-helix  
 FT Region 41..242  
 FT /label= Repeat 1  
 FT /notes= "alpha-helical repeat region"  
 FT Region 69..637  
 FT /label= Coiled-coil  
 FT /notes= "breaks in 7-residue periodicity of the coiled-coil occur at amino acids 136-141, 261-304 and 383-387"  
 FT Region 332..492  
 FT /label= Repeat 2  
 FT /notes= "alpha-helical repeat region"  
 FT Domain 627..689  
 FT /label= Proline-rich  
 FT Domain 913..929  
 FT /label= C-terminal

WO9709994-A1.  
 XX  
 XX 20-MAR-1997.  
 XX  
 XX 16-SEP-1996; 96WO-US014819.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 XX WPI; 1997-202002/18.  
 DR N-PSDB; AAT61728.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 XX Disclosure; Fig 13; 296pp; English.

XX This sequence comprises the pneumococcal protein surface C (pspC) of  
 CC Streptococcus pneumoniae strain EF6796. The sequence was deduced from the  
 CC pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil  
 CC region, proline-rich central region, repeat regions, with a choline  
 CC binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides

CC share 3 regions of high sequence identity. One is a protection-eliciting  
 CC region present within the alpha-helical domain. The others are the  
 CC proline-rich domain and a repeat domain shared with other choline-binding  
 CC proteins and thought to play a role in cell surface association. PspC and  
 CC PspA polypeptides, and their fragments, can be used in vaccines to  
 CC protect against S. pneumoniae infection, and hence for the prevention of  
 CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.  
 CC (Updated on 17-OCT-2003 to standardise OS field)  
 XX

SQ Sequence 929 AA;

Query Match 93.0%; Score 459.5; DB 2; Length 929;  
 Best Local Similarity 96.0%; Pred. No. 8.7e-35;  
 Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKGFRAPLOSQKIDAKKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
 |||||  
 Db 530 LKEIDSDSDYAKGFRAPLOSQKIDAKKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 588  
 |||||  
 QY 61 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAADLKKAVDEPE 100  
 |||||  
 Db 589 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAADLKKAVDEPE 628  
 |||||

## RESULT 5

AAV43384  
 ID AAV43384 standard; protein; 929 AA.  
 XX

AC AAV43384;

DT 27-JAN-2000 (first entry)

DE S. pneumoniae PspC protein sequence.

XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;  
 KW epitopic region; immunogenic composition; vaccine composition; therapy;  
 KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;  
 KW anti-PspA antibody.

XX Streptococcus pneumoniae.

XX WO9953940-A1.

XX 28-OCT-1999.

XX 23-APR-1999; 99WO-US008895.

XX 23-APR-1998; 98US-0082728P.

XX (UYAL-) UNIV ALABAMA.

XX Briles DE, Hollingshead SK, Brooks-Walter A;

XX WPI; 1999-620581/53.

XX N-PSDB; AAZ31956.

XX New epitope polypeptides of Pneumococcal surface protein C, used to  
 PT develop products for immunological, immunogenic or vaccine compositions,  
 PT particularly against Streptococcus pneumoniae infections.

XX Example; Fig 11; 109pp; English.

XX This sequence is the Streptococcus pneumoniae pneumococcal surface  
 CC protein C. The invention relates to an isolated and/or purified  
 CC polypeptide comprising at least one epitope or epitopic region of  
 CC Pneumococcal surface protein C (PspC). The polypeptides or vectors  
 CC containing sequence encoding them can be used as immunogenic,  
 CC immunological or vaccine compositions. The compositions can be used for  
 CC eliciting an immunological response against Streptococcus pneumoniae  
 CC (SP), which can cause otitis media, meningitis, bacteraemia and  
 CC pneumonia. They can be used for eliciting an anti-PspA antibody. The  
 CC nucleic acid molecules can also be used for detecting pspC, pspA or SP

XX

```
SQ Sequence 929 AA;
Query Match 93.0%; Score 459.5; DB 2; Length 929;
Best Local Similarity 96.0%; Pred. No. 8.7e-35;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
|||||
Db 3667 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 3725
|||||

Qy 61 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100
|||||
Db 3726 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 3765
|||||

RESULT 7
AAW14580
ID AAW14580 standard; protein; 188 AA.
XX
AC AAW14580;
XX
XX 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
DE
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteriaemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Rct135.
OS
XX
XX WO9709994-A1.
DN
XX
XX 20-MAR-1997.
PD
XX
XX 16-SEP-1996; 96WO-US014819.
PF
XX
XX 15-SEP-1995; 95US-00529055.
PR
XX
XX (UABR-) UAB RES FOUND.
PA
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
DR
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
PT
XX
XX Example 6; Fig 13; 296pp; English.
PS
XX
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologues. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 188 AA;

Query Match 92.4%; Score 456.5; DB 2; Length 188;
Best Local Similarity 95.0%; Pred. No. 2.5e-35;
Matches 95; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
|||||
Db 1 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 59
|||||

Qy 61 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100
|||||
Db 60 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99
|||||
```

```
SQ Sequence 929 AA;
Query Match 93.0%; Score 459.5; DB 2; Length 929;
Best Local Similarity 96.0%; Pred. No. 8.7e-35;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
|||||
Db 530 LKEIDESDSEYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 588
|||||

Qy 61 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100
|||||
Db 589 EGNNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628
|||||

RESULT 6
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
AC ABU08487;
XX
XX 24-JUN-2003 (first entry)
DT
XX
XX S. pneumoniae pneumococcal surface protein A (PspA) protein.
DE
XX
XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
XX Streptococcus pneumoniae.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 1..8991
FT /note= "All Xaa residues within this sequence are
FT unknown"
XX
XX US6500613-B1.
PN
XX
XX 31-DEC-2002.
PD
XX
XX 16-SEP-1996; 96US-00714741.
PF
XX
XX 15-SEP-1995; 95US-00529055.
PR
XX
XX (UVAL-) UNIV ALABAMA.
PA
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
DR
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
PT
XX
XX Disclosure; Col 145-188; 186pp; English.
PS
XX
XX The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA) -
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents S. pneumoniae
CC PspA protein
XX
SQ Sequence 8991 AA;

Query Match 93.0%; Score 459.5; DB 6; Length 8991;
Best Local Similarity 96.0%; Pred. No. 1.3e-33;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```



```
RESULT 8
ABW02613
ID ABW02613 standard; protein; 188 AA.
XX
AC ABW02613;
XX
DT 12-FEB-2004 (first entry)
XX
XX Rct129c pneumococcal surface protein A (PspA) central region.
DE
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
PI WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 59; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Rct129c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX Sequence 188 AA;
SQ
Query Match 92.4%; Score 456.5; DB 7; Length 188;
Best Local Similarity 95.0%; Pred. No. 2.5e-35;
Matches 95; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 1 LKXEDSDSDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60
DB 1 LKXEDSDSDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLE-VOLKDA 59
QY 61 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 100
DB 60 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPD 99
RESULT 9
ABU08491
ID ABU08491 standard; protein; 588 AA.
XX
XX ABU08491;
AC
```

```
XX
DT 24-JUN-2003 (first entry)
XX
XX Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.
XX
XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
XX Streptococcus pneumoniae.
XX
XX US6500613-B1.
XX
XX 31-DEC-2002.
XX
XX 16-SEP-1996; 96US-00714741.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
XX Disclosure; Fig 23; 186pp; English.
XX
XX The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC) and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents a coiled coil
CC motif of the alpha-helix of S. pneumoniae PspC protein
XX
XX Sequence 588 AA;
SQ
Query Match 90.6%; Score 447.5; DB 6; Length 588;
Best Local Similarity 95.9%; Pred. No. 6.9e-34;
Matches 94; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 LKXEDSDSDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60
DB 492 LKXEDSDSDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLE-VOLKDA 550
QY 61 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDE 98
DB 551 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAVDE 588
RESULT 10
AA43392
ID AA43392 standard; protein; 589 AA.
XX
XX AA43392;
AC
XX
XX 27-JAN-2000 (first entry)
DT
XX
XX PspC alpha-helix coiled-coil region.
DE
XX
XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
XX Streptococcus pneumoniae.
```

```

XX WO9953940-A1.
PN
XX
XX 28-OCT-1999.
PD
XX
XX 23-APR-1999; 99WO-US008895.
PF
XX
XX 23-APR-1998; 98US-0082728P.
PR
XX
XX (UYAL-) UNIV ALABAMA.
PA
XX
XX Briles DE, Hollingshead SK, Brooks-Walter A;
PI
XX WPI; 1999-620581/53.
XX
XX New epitope polypeptides of Pneumococcal surface protein C, used to
PT develop products for immunological, immunogenic or vaccine compositions,
PT particularly against Streptococcus pneumoniae infections.
XX
XX Example 1; Fig 3; 109pp; English.
PS
XX This sequence is the coiled-coil region of the Streptococcus pneumoniae
CC pneumococcal surface protein C. The invention relates to an isolated
CC and/or purified polypeptide comprising at least one epitope or epitopic
CC region of Pneumococcal surface protein C (PspC). The polypeptides or
CC vectors containing sequence encoding them can be used as immunogenic,
CC immunological or vaccine compositions. The compositions can be used for
CC eliciting an immunological response against Streptococcus pneumoniae
CC (SP), which can cause otitis media, meningitis, bacteraemia and
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX
XX Sequence 589 AA;
SQ
Query Match 90.6%; Score 447.5; DB 2; Length 589;
Best Local Similarity 95.9%; Pred. No. 7e-34;
Matches 94; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAIAKLECVQLKDA 60
DB 493 LKEIDSESDYLYKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAIAKLE-VQLKDA 551
QY 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDE 98
DB 552 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDE 589
RESULT 11
AAW14578
ID AAW14578 standard; protein; 204 AA.
XX
XX AAW14578;
AC
XX
XX 17-OCT-2003 (revised)
DT
XX 28-OCT-1997 (first entry)
DT
XX Streptococcus pneumoniae PspA central region.
DE
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
KW
XX Streptococcus pneumoniae; strain Rct123.
OS
XX
XX Key Location/Qualifiers
FH
XX Misc-difference 4
FT
XX /note= "unidentified amino acid"
FT
XX Misc-difference 8
FT
XX /note= "unidentified amino acid"
FT
XX
XX WO9709994-A1.
PN
XX 20-MAR-1997.
PD
XX
XX

```

```

PF 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
PR
XX (UABR-) UAB RES FOUND.
PA
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
PS
XX This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 204 AA;
SQ
Query Match 90.2%; Score 445.5; DB 2; Length 204;
Best Local Similarity 93.0%; Pred. No. 3.1e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAIAKLECVQLKDA 60
DB 1 IKEDEXSSEDYLYKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAIAKLE-VQLKDA 59
QY 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDE 100
DB 60 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVIDE 99
RESULT 12
ABW02612
ID ABW02612 standard; protein; 204 AA.
XX
XX ABW02612;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX Rct123c pneumococcal surface protein A (PspA) central region.
DE
XX
XX pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
KW
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH
XX Misc-difference 1..204
FT
XX /note= "Xaa = Unknown amino acid"
FT
XX
XX US6592876-B1.
PN
XX 15-JUL-2003.
PD
XX
XX 15-SEP-1995; 95US-00529055.
PF
XX 20-APR-1993; 93US-00048896.
PR
XX 06-JUN-1995; 95US-00465746.
PD
XX
XX

```

PA (UABR-) UAB RES FOUND.  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.  
XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding  
PT pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 58; 121pp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
CC at least 2 different full length isolated genes encoding pneumococcal  
CC surface protein A (PspAs) from different groups based on restriction  
CC fragment polymorphism analysis. The invention is useful for obtaining  
CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Rct123c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
CC used in the exemplification of the invention  
XX  
XX Sequence 204 AA;  
SQ  
Query Match 90.2%; Score 445.5; DB 7; Length 204;  
Best Local Similarity 93.0%; Pred. No. 3.1e-34;  
Matches 93; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
QY 1 LKEIDESDSEYAKGFRAPLOSGLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
Db :|||||  
1 IXEXDSXSEDYKGEFRAPLOSGLDKAKKLSKLEELSDKIDELDAETAKLE-VOLKDA 59  
QY 61 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 100  
Db :|||||  
60 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99  
RESULT 13  
AAW14562  
ID AAW14562 standard; protein; 180 AA.  
XX  
XX AAW14562;  
AC  
XX  
XX 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
XX Streptococcus pneumoniae PspA central region.  
DE  
XX  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
XX Streptococcus pneumoniae; strain 0922134c.  
OS  
XX  
XX WO9709994-A1.  
PN  
XX  
XX 20-MAR-1997.  
PD  
XX  
XX 16-SEP-1996; 96WO-US014819.  
PF  
XX  
XX 15-SEP-1995; 95US-00529055.  
PR  
XX  
XX (UABR-) UAB RES FOUND.  
PA  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX  
XX Example 6; Fig 13; 296pp; English.  
XX  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain 0922134c.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
XX Sequence 180 AA;  
SQ  
Query Match 90.1%; Score 445; DB 2; Length 180;  
Best Local Similarity 95.0%; Pred. No. 3e-34;  
Matches 95; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
QY 1 LKEIDESDSEYAKGFRAPLOSGLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
Db :|||||  
1 LKEIDESDSEYKLEGLRAPLOSGLDKAKKLSKLEELSDKIDELDAETAKLE-VOLKDA 59  
QY 61 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 100  
Db :|||||  
60 EGNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 98  
RESULT 14  
AAW14579  
ID AAW14579 standard; protein; 187 AA.  
XX  
XX AAW14579;  
AC  
XX  
XX 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX  
XX Streptococcus pneumoniae PspA central region.  
DE  
XX  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX  
XX Streptococcus pneumoniae; strain Rct129.  
OS  
XX  
XX WO9709994-A1.  
PN  
XX  
XX 20-MAR-1997.  
PD  
XX  
XX 16-SEP-1996; 96WO-US014819.  
PF  
XX  
XX 15-SEP-1995; 95US-00529055.  
PR  
XX  
XX (UABR-) UAB RES FOUND.  
PA  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
XX WPI; 1997-202002/18.  
DR  
XX  
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX  
XX Example 6; Fig 13; 296pp; English.  
XX  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct129.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against *S. pneumoniae* infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 187 AA;

Query Match 89.5%; Score 442; DB 2; Length 187;  
Best Local Similarity 94.0%; Pred. No. 6e-34;  
Matches 94; Conservative 1; Mismatches 3; Indels 2; Gaps 2;  
QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDA 60  
DB 1 LKEIDSESDYAKGFRAPLQSKLDYTKAKLSKLEELSDKIDELDAEIAKLE-VQLKDA 59  
QY 61 EGNNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 100  
DB 60 EGNNNVEAYFKEGLEKTTAEKKAELEKAE-DLKKAVIDEPE 98

RESULT 15  
ABU08490  
ID ABU08490 standard; protein; 1231 AA.  
XX  
AC ABU08490;  
XX  
DT 24-JUN-2003 (first entry)  
XX  
DE Fragment of *S. pneumoniae* PspC protein.  
XX  
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
KW antibacterial.  
XX  
OS Streptococcus pneumoniae.  
XX  
FN US6500613-B1.  
XX  
PD 31-DEC-2002.  
XX  
PF 16-SEP-1996; 96US-00714741.  
XX  
PR 15-SEP-1995; 95US-00529055.  
XX  
PA (UYAL-) UNIV ALABAMA.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX  
DR WPI; 2003-361534/34.  
XX

Isolated PspC amino acid sequence used as polymerase chain reaction or  
hybridization probe, comprises pneumococcal surface protein having alpha-  
helical, proline rich and repeat regions.

Disclosure; Fig 22; 186pp; English.

The present invention relates to the isolation of Streptococcus  
pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
like protein having alpha-helical, proline rich and repeat regions. The  
PspC and PspA proteins may be used in a vaccine to protect against  
pneumococcal infections. The polynucleotide sequences encoding PspC and  
PspA may be used for the expression of the proteins, and as PCR primers  
or hybridisation probes. The present sequence represents a fragment of *S.*  
pneumoniae PspC protein

XX  
SQ Sequence 1231 AA;  
Query Match 89.0%; Score 439.5; DB 6; Length 1231;  
Best Local Similarity 93.9%; Pred. No. 9.5e-33;  
Matches 92; Conservative 2; Mismatches 3; Indels 1; Gaps 1;  
QY 3 EIDESDSEYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDAEG 62  
DB 494 EVQLSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLE-DQLKDAEG 552  
QY 63 NNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 100  
DB 553 NNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 590  
Search completed: June 18, 2005, 16:51:23  
Job time : 73.0731 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:36:26 ; Search time 18.9189 Seconds  
(without alignments)  
394.574 Million cell updates/sec

Title: US-10-674-755-10  
Perfect score: 494  
Sequence: 1 LKEIDSESDYAKGFRAP.....KKAELKAEADLKKAVDEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	100.0	100	4	US-09-147-875A-10
2	464.5	94.0	99	2	US-08-710-749-17
3	459.5	93.0	170	4	US-08-529-055-60
4	459.5	93.0	181	4	US-08-529-055-42
5	459.5	93.0	864	4	US-08-714-741-40
6	459.5	93.0	8991	4	US-08-714-741-32
7	456.5	92.4	99	4	US-09-147-875A-16
8	456.5	92.4	188	4	US-08-529-055-59
9	447.5	90.6	141	4	US-09-286-981B-2
10	447.5	90.6	588	4	US-08-714-741-42
11	445.5	90.2	99	2	US-08-710-749-15
12	445.5	90.2	204	4	US-08-529-055-58
13	444.5	90.0	99	4	US-09-147-875A-15
14	439.5	89.0	1231	4	US-08-714-741-41
15	432.5	87.6	206	4	US-08-529-055-54
16	430.5	87.1	99	2	US-08-710-749-14
17	429.5	86.9	99	4	US-09-147-875A-14
18	420.5	85.1	99	2	US-08-710-749-11
19	420.5	85.1	198	4	US-08-529-055-61
20	420.5	85.1	619	1	US-08-465-746-2
21	420.5	85.1	619	1	US-08-214-164-2
22	420.5	85.1	619	2	US-08-467-852A-3
23	420.5	85.1	619	2	US-08-246-636-2
24	420.5	85.1	619	2	US-08-247-491A-3
25	420.5	85.1	619	2	US-08-319-795-2
26	420.5	85.1	619	2	US-08-468-985-2
27	420.5	85.1	619	3	US-08-312-949-2

```

28 420.5 85.1 648 1 US-08-072-070-2 Sequence 2, Appli
29 420.5 85.1 648 1 US-08-469-434-2 Sequence 2, Appli
30 420.5 85.1 648 1 US-08-214-222-2 Sequence 2, Appli
31 420.5 85.1 648 2 US-08-467-852A-2 Sequence 2, Appli
32 420.5 85.1 648 2 US-08-468-718-2 Sequence 2, Appli
33 420.5 85.1 648 2 US-08-247-491A-2 Sequence 2, Appli
34 420.5 85.1 648 3 US-08-446-201-3 Sequence 3, Appli
35 420.5 85.1 695 1 US-08-127-499A-23 Sequence 23, Appli
36 420.5 85.1 695 1 US-08-482-847-23 Sequence 13, Appli
37 418.5 84.7 99 2 US-08-710-749-13 Sequence 13, Appli
38 412.5 83.5 99 2 US-08-710-749-10 Sequence 10, Appli
39 412.5 83.5 99 4 US-09-147-875A-11 Sequence 11, Appli
40 412.5 83.5 204 4 US-08-529-055-51 Sequence 51, Appli
41 412 83.4 100 4 US-09-147-875A-12 Sequence 12, Appli
42 408.5 82.7 288 3 US-08-312-949-4 Sequence 4, Appli
43 408.5 82.7 288 3 US-08-446-201-4 Sequence 4, Appli
44 398 80.6 289 1 US-08-072-070-4 Sequence 4, Appli
45 398 80.6 289 1 US-08-469-434-4 Sequence 4, Appli

```

## ALIGNMENTS

```

RESULT 1
US-09-147-875A-10
; Sequence 10, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-10

```

Query Match 100.0%; Score 494; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.6e-41;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60

Qy 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100
Db 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100

```

```

RESULT 2
US-08-710-749-17
; Sequence 17, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:

```

```
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/710,749
;; FILING DATE: 20-SEP-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer, William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454312-2074
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; INFORMATION FOR SEQ ID NO: 17:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 99 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: n/a
;; TOPOLOGY: linear
;; MOLECULE TYPE: amino acid
US-08-710-749-17

Query Match 94.0%; Score 464.5; DB 2; Length 99;
Best Local Similarity 97.0%; Pred. No. 1.3e-38;
Matches 97; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLE-VQLKDA 59

Qy 61 EGNNNVEAYFKGLEKTTAAEKAELEKAEADLKAVDEPE 100
Db 60 EGNNNVEAYFKGLEKTTAAEKATELEKAEADLKRAVDEPE 99

RESULT 3
US-08-529-055-60
; Sequence 60, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-42

Query Match 93.0%; Score 459.5; DB 4; Length 170;
Best Local Similarity 96.0%; Pred. No. 7.6e-38;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLE-VQLKDA 59

Qy 61 EGNNNVEAYFKGLEKTTAAEKAELEKAEADLKAVDEPE 100
Db 60 EGNNNVEAYFKGLEKTTAAEKAELEKAEADLKRAVDEPE 99

RESULT 4
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-42

Query Match 93.0%; Score 459.5; DB 4; Length 181;
Best Local Similarity 96.0%; Pred. No. 8.2e-38;
```

Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVOLKDA 60  
 Db 1 LKEIDESDSYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 59  
 QY 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100  
 Db 60 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 5

US-08-714-741-40  
 ; Sequence 40, Application US/08714741  
 ; Patent No. 6500613  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: McDaniel, Larry S.  
 ; APPLICANT: Swiatlo, Edwin  
 ; APPLICANT: Yother, Janet  
 ; APPLICANT: Crain, Marilyn J.  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Tart, Rebecca  
 ; APPLICANT: Brooks-Walter, Alexis  
 ; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
 ; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
 ; NUMBER OF SEQUENCES: 47  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford, P.C.  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION NUMBER: US/08/714,741  
 ; FILING DATE: 16-SEP-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer Esq., William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454312-2460  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 40:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 864 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: amino acid  
 ; US-08-714-741-40

Query Match 93.0%; Score 459.5; DB 4; Length 864;  
 Best Local Similarity 96.0%; Pred. No. 5.3e-37;  
 Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVOLKDA 60  
 Db 465 LKEIDESDSYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 523  
 QY 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100  
 Db 524 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 563

RESULT 6

US-08-714-741-32  
 ; Sequence 32, Application US/08714741  
 ; Patent No. 6500613  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: McDaniel, Larry S.  
 ; APPLICANT: Swiatlo, Edwin  
 ; APPLICANT: Yother, Janet  
 ; APPLICANT: Crain, Marilyn J.  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Tart, Rebecca  
 ; APPLICANT: Brooks-Walter, Alexis  
 ; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
 ; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
 ; NUMBER OF SEQUENCES: 47  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford, P.C.  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION NUMBER: US/08/714,741  
 ; FILING DATE: 16-SEP-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer Esq., William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454312-2460  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8991 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: amino acid  
 ; US-08-714-741-32

Query Match 93.0%; Score 459.5; DB 4; Length 8991;  
 Best Local Similarity 96.0%; Pred. No. 8.6e-36;  
 Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 LKEIDESDSYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVOLKDA 60  
 Db 3667 LKEIDESDSYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 3725  
 QY 61 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100  
 Db 3726 EGNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 3765

RESULT 7

US-09-147-875A-16  
 ; Sequence 16, Application US/09147875A  
 ; Patent No. 6638516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BECKER et al.  
 ; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
 ; FILE REFERENCE: 454312-2471  
 ; CURRENT APPLICATION NUMBER: US/09/147,875A  
 ; CURRENT FILING DATE: 1999-05-24  
 ; NUMBER OF SEQ ID NOS: 28

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-16

Query Match      92.4%; Score 456.5; DB 4; Length 99;
Best Local Similarity 96.0%; Pred. No. 7.8e-38;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 100
    |||||
Db 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 100
    |||||

Qy 61 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 100
    |||||
Db 60 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 99
    |||||

RESULT 8
US-08-529-055-59
; Sequence 59, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-59

Query Match      92.4%; Score 456.5; DB 4; Length 188;
Best Local Similarity 95.0%; Pred. No. 3.7e-37;
Matches 95; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 60
    |||||
Db 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 60
    |||||

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-16

Query Match      92.4%; Score 456.5; DB 4; Length 99;
Best Local Similarity 96.0%; Pred. No. 7.8e-38;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 100
    |||||
Db 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 100
    |||||

Qy 61 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 100
    |||||
Db 60 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 99
    |||||

RESULT 9
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wisemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-2

Query Match      90.6%; Score 447.5; DB 4; Length 141;
Best Local Similarity 95.9%; Pred. No. 9.2e-37;
Matches 94; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 60
    |||||
Db 45 LKEIDESDSEYAKGFRAPLQSKLDKAKKAELEKAEADLKKAADPE 103
    |||||

Qy 61 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 98
    |||||
Db 104 EGNNNVEAYFKGLEKTTAEKKAELKAEADLKKAADPE 141
    |||||

RESULT 10
US-08-714-741-42
; Sequence 42, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
```



```
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-42

Query Match          90.6%; Score 447.5; DB 4; Length 588;
Best Local Similarity 95.9%; Pred. No. 5e-36;
Matches 94; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKGFRAPLQSKLDKAKKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
Db 492 LKEIDSDSDYKKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 550

QY 61 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDE 98
Db 551 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDE 588

RESULT 11
US-08-710-749-15
; Sequence 15, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-15
```

```
Query Match          90.2%; Score 445.5; DB 2; Length 99;
Best Local Similarity 93.0%; Pred. No. 9.4e-37;
Matches 93; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKGFRAPLQSKLDKAKKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
Db 1 LKEIDSDSDYKKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 59

QY 61 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDEPE 100
Db 60 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDEPE 99

RESULT 12
US-08-529-055-58
; Sequence 58, Application US/08529055
; Patent No. 652876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-58

Query Match          90.2%; Score 445.5; DB 4; Length 204;
Best Local Similarity 93.0%; Pred. No. 2.2e-36;
Matches 93; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKGFRAPLQSKLDKAKKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
Db 1 IKEXDESXSDYKKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 59

QY 61 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDEPE 100
Db 60 EGNNNVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDEPE 99
```

RESULT 13  
US-09-147-875A-15  
; Sequence 15, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-15  
  
Query Match 90.0%; Score 444.5; DB 4; Length 99;  
Best Local Similarity 93.0%; Pred. No. 1.2e-36;  
Matches 93; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
  
Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLECVQLKDA 60  
Db 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAELE-VQLKDA 59  
  
Qy 61 EGNNNVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 100  
Db 60 EGNNNVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99  
  
RESULT 14  
US-08-714-741-41  
; Sequence 41, Application US/08714741  
; Patent No. 6500613  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
; TITLE OF INVENTION: PORTIONS AND PRODUCTS  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,741  
; FILING DATE: 16-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer Esq., William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2460  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 41:  
; 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLECVQLKDA 60

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1231 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-714-741-41  
  
Query Match 89.0%; Score 439.5; DB 4; Length 1231;  
Best Local Similarity 93.9%; Pred. No. 7.4e-35;  
Matches 92; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
  
Qy 3 BIDESDSEYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLECVQLKDAEG 62  
Db 494 EVQLSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEB-DQLKDAEG 552  
  
Qy 63 NNNVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 100  
Db 553 NNNVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 590  
  
RESULT 15  
US-08-529-055-54  
; Sequence 54, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-54  
  
Query Match 87.6%; Score 432.5; DB 4; Length 206;  
Best Local Similarity 90.0%; Pred. No. 4.3e-35;  
Matches 90; Conservative 4; Mismatches 5; Indels 1; Gaps 1;  
  
Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLECVQLKDA 60

Search completed: June 18, 2005, 17:07:08  
Job time : 18.9189 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 17:01:47 ; Search time 62.963 Seconds  
(without alignments)  
609.850 Million cell updates/sec

Title: US-10-674-755-10

Perfect score: 494

Sequence: 1 LKEIDSDSDYAKGFRAP.....KKALEKAEADLKKAVDEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	100.0	100	15	US-10-674-755-10
2	459.5	93.0	170	15	US-10-299-636-75
3	459.5	93.0	181	15	US-10-299-636-57
4	459.5	93.0	643	15	US-10-299-636-95
5	459.5	93.0	670	9	US-09-748-875-63
6	459.5	93.0	670	10	US-09-298-523B-63
7	459.5	93.0	690	9	US-09-748-875-61
8	459.5	93.0	690	10	US-09-298-523B-61
9	459.5	93.0	691	9	US-09-748-875-1
10	459.5	93.0	691	10	US-09-298-523B-1
11	459.5	93.0	701	9	US-09-748-875-62

12	459.5	93.0	701	10	US-09-298-523B-62	Sequence 62, Appl
13	459.5	93.0	707	9	US-09-748-875-2	Sequence 2, Appl
14	459.5	93.0	707	10	US-09-298-523B-2	Sequence 2, Appl
15	459.5	93.0	711	9	US-09-748-875-3	Sequence 3, Appl
16	459.5	93.0	711	10	US-09-298-523B-3	Sequence 3, Appl
17	459.5	93.0	739	17	US-10-732-923-3294	Sequence 3294, Ap
18	459.5	93.0	929	9	US-09-748-875-60	Sequence 60, Appl
19	459.5	93.0	929	10	US-09-298-523B-60	Sequence 60, Appl
20	459.5	93.0	929	15	US-10-299-636-94	Sequence 94, Appl
21	456.5	92.4	99	15	US-10-674-755-16	Sequence 16, Appl
22	456.5	92.4	188	15	US-10-299-636-74	Sequence 74, Appl
23	447.5	90.6	141	14	US-10-254-995-2	Sequence 2, Appl
24	447.5	90.6	589	9	US-09-748-875-14	Sequence 14, Appl
25	447.5	90.6	589	10	US-09-298-523B-14	Sequence 14, Appl
26	447.5	90.6	589	15	US-10-299-636-97	Sequence 97, Appl
27	445.5	90.2	204	15	US-10-299-636-73	Sequence 73, Appl
28	444.5	90.0	99	15	US-10-674-755-15	Sequence 15, Appl
29	432.5	87.6	206	15	US-10-299-636-69	Sequence 69, Appl
30	429.5	86.9	99	15	US-10-674-755-14	Sequence 14, Appl
31	420.5	85.1	198	15	US-10-299-636-105	Sequence 105, App
32	420.5	85.1	354	15	US-10-299-636-96	Sequence 96, Appl
33	420.5	85.1	588	15	US-10-299-636-96	Sequence 96, Appl
34	420.5	85.1	619	10	US-09-882-774-1	Sequence 1, Appl
35	420.5	85.1	619	15	US-10-282-122A-73702	Sequence 73702, A
36	420.5	85.1	619	16	US-10-414-532-72	Sequence 72, Appl
37	412.5	83.5	99	15	US-10-674-755-11	Sequence 11, Appl
38	412.5	83.5	204	15	US-10-299-636-66	Sequence 66, Appl
39	412	83.4	100	15	US-10-674-755-12	Sequence 12, Appl
40	393.5	79.7	195	15	US-10-299-636-86	Sequence 86, Appl
41	389.5	78.8	99	15	US-10-674-755-13	Sequence 13, Appl
42	363	73.5	336	15	US-10-299-636-103	Sequence 103, Appl
43	346.5	70.1	193	15	US-10-299-636-64	Sequence 64, Appl
44	343.5	69.5	99	15	US-10-674-755-17	Sequence 17, Appl
45	328	66.4	183	15	US-10-299-636-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-10-674-755-10

Sequence 10, Application US/10674755

Publication No. US20040067237A1

GENERAL INFORMATION:

APPLICANT: BECKER et al.

TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

FILE REFERENCE: 454312-2471

CURRENT APPLICATION NUMBER: US/10/674,755

CURRENT FILING DATE: 2003-09-30

PRIOR FILING DATE: US/09/147,875A

PRIOR FILING DATE: 1999-05-24

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 100

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

US-10-674-755-10

Query Match 100.0%; Score 494; DB 15; Length 100;

Best Local Similarity 100.0%; Pred. No. 3.2e-35;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAETAKLECVQLKDA 60

Db 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSDKIDELDAETAKLECVQLKDA 60

Qy 61 EGNNVNVEAFKEGLEKTTAKKAELEKAEADLKKAVDEPE 100

Db 61 EGNNVNVEAFKEGLEKTTAKKAELEKAEADLKKAVDEPE 100

RESULT 2

```
US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75
Query Match 93.0%; Score 459.5; DB 15; Length 170;
Best Local Similarity 96.0%; Pred. No. 5.4e-32;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVOLKDA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 59

Qy 61 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 100
Db 60 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 99

RESULT 3
US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57
Query Match 93.0%; Score 459.5; DB 15; Length 181;
Best Local Similarity 96.0%; Pred. No. 5.8e-32;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVOLKDA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 59

Qy 61 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 100
Db 60 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 99

RESULT 4
US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95
Query Match 93.0%; Score 459.5; DB 15; Length 643;
Best Local Similarity 96.0%; Pred. No. 2.4e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAEIAKLECVOLKDA 60
Db 245 LKEIDSESDYAKGFRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLE-VOLKDA 303

Qy 61 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 100
Db 304 EGNNNVEAYFKGLEKTTAAEKKAELKAEADLKKAVDPE 343

RESULT 5
US-09-748-875-63
; Sequence 63, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-63
```

```
Query Match      93.0%; Score 459.5; DB 9; Length 670;
Best Local Similarity 96.0%; Pred. No. 2.5e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
    |||||
Db 498 LKEIDESDSDYKLEGRAPLQSKLDTTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 556

QY 61 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 100
    |||||
Db 557 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 596

RESULT 6
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-63

Query Match      93.0%; Score 459.5; DB 10; Length 670;
Best Local Similarity 96.0%; Pred. No. 2.5e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
    |||||
Db 498 LKEIDESDSDYKLEGRAPLQSKLDTTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 556

QY 61 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 100
    |||||
Db 557 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 596

RESULT 7
US-09-748-875-61
; Sequence 61, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match      93.0%; Score 459.5; DB 9; Length 690;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
    |||||
Db 529 LKEIDESDSDYKLEGRAPLQSKLDTTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 587
```

```
QY 61 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 100
    |||||
Db 588 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 627

RESULT 8
US-09-298-523B-61
; Sequence 61, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-61

Query Match      93.0%; Score 459.5; DB 10; Length 690;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
    |||||
Db 529 LKEIDESDSDYKLEGRAPLQSKLDTTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 587

QY 61 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 100
    |||||
Db 588 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 627

RESULT 9
US-09-748-875-1
; Sequence 1, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-1

Query Match      93.0%; Score 459.5; DB 9; Length 691;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLSKLEELSDKIDELDAETAKLECVOLKDA 60
    |||||
Db 530 LKEIDESDSDYKLEGRAPLQSKLDTTKAKLSKLEELSDKIDELDAETAKLE-VOLKDA 588

QY 61 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 100
    |||||
Db 589 EGNNNVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 628

RESULT 10
```

```
US-09-298-523B-1
; Sequence 1, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-1
Query Match          93.0%; Score 459.5; DB 10; Length 691;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 530 LKEIDSESDYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VQLKDA 588
Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 100
Db 589 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 628
RESULT 11
US-09-748-875-62
; Sequence 62, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-62
Query Match          93.0%; Score 459.5; DB 9; Length 701;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 529 LKEIDSESDYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VQLKDA 587
Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 100
Db 588 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 627
RESULT 12
US-09-298-523B-62
; Sequence 62, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
```

```
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-62
Query Match          93.0%; Score 459.5; DB 10; Length 701;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 529 LKEIDSESDYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VQLKDA 587
Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 100
Db 588 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 627
RESULT 13
US-09-748-875-2
; Sequence 2, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-2
Query Match          93.0%; Score 459.5; DB 9; Length 707;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLEELSDKIDELDAEIAKLECVQLKDA 60
Db 530 LKEIDSESDYAKGFRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAKLE-VQLKDA 588
Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 100
Db 589 EGNNVVAYFKGLEKTTAEKKAELKAEADLKKAVDPE 628
RESULT 14
US-09-298-523B-2
; Sequence 2, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 707
```



```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-2

Query Match      93.0%; Score 459.5; DB 10; Length 707;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKEGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
    |||||
Db 530 LKEIDSDSDYAKEGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
    |||||

QY 61 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 100
    |||||
Db 589 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 628
    |||||

RESULT 15
US-09-748-875-3
; Sequence 3, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-3

Query Match      93.0%; Score 459.5; DB 9; Length 711;
Best Local Similarity 96.0%; Pred. No. 2.6e-31;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYAKEGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
    |||||
Db 539 LKEIDSDSDYAKEGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
    |||||

QY 61 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 100
    |||||
Db 598 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 637
    |||||
```

Search completed: June 18, 2005, 18:00:24  
Job time : 62.963 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2005, 16:35:36 ; Search time 13.013 Seconds  
(without alignments)  
739.389 Million cell updates/sec

Title: US-10-674-755-10  
Perfect score: 494  
Sequence: 1 LKEIDSESDYAKGFRAP.....KKALEKAEADLKKAVDPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.5	85.1	619	2 A97887	surface protein ps
2	420.5	85.1	619	2 A41971	surface protein ps
3	114	23.1	744	2 F95013	pneumococcal surfa
4	113	22.9	161	2 S48396	tropomyosin TPM2 -
5	106	21.5	852	2 D72230	conserved hypotet
6	103	20.9	233	2 S70531	bbk2.11 protein pr
7	102	20.6	3488	2 T34418	hypothetical prote
8	100	20.2	229	2 S70532	outer surface prot
9	99	20.0	1006	2 C70445	ATPase subunit of
10	99	20.0	1818	1 S73852	hypothetical prote
11	97.5	19.7	395	2 AC1754	capsid protein lba
12	96.5	19.5	1410	1 A57013	early endosome ant
13	95.5	19.3	388	2 A46173	Mrp4 protein - Str
14	95.5	19.3	405	2 A33939	Fc gamma (IGG) rec
15	95.5	19.3	764	2 T05409	hypothetical prote
16	95.5	19.3	1177	2 B75150	chromosome segrega
17	94.5	19.1	369	2 AG1648	hypothetical prote
18	94	19.0	284	2 S58921	tropomyosin isofo
19	94	19.0	284	2 S58922	tropomyosin isofo
20	93.5	18.9	399	2 E71169	hypothetical prote
21	93.5	18.9	1319	2 A28313	glued protein - fr
22	93	18.8	284	2 A44980	tropomyosin, obliq
23	93	18.8	1110	2 IS1116	NP-180 - sea lampr
24	92.5	18.7	1156	2 B70356	chromosome assembl
25	92.5	18.7	2116	2 A26655	myosin heavy chain
26	92	18.6	431	2 A97225	KBP-type peptidyl
27	92	18.6	2401	2 T28676	rhoptyr protein -
28	91.5	18.5	886	2 H69378	conserved hypotet
29	91.5	18.5	1169	2 A64505	p115 homolog - Met

30	91.5	18.5	1938	1 JX0178	myosin heavy chain
31	91	18.4	879	2 C71083	conserved hypotet
32	91	18.4	880	2 F75103	conserved hypotet
33	91	18.4	2139	2 T18296	myosin heavy chain
34	90.5	18.3	387	2 S57834	fcra protein precu
35	90.5	18.3	433	2 A89951	trigger factor [lm
36	90.5	18.3	1078	2 T18352	protein p120 - Myc
37	90.5	18.3	1790	2 S67593	transport protein f
38	90	18.2	896	2 S43074	epidermal growth f
39	90	18.2	1133	2 T32976	hypothetical prote
40	90	18.2	1312	2 T30845	probable DNA repai
41	89.5	18.1	258	2 A02985	myosin heavy chain
42	89.5	18.1	421	2 JV0057	to1a protein - Bsc
43	89.5	18.1	520	2 S35575	myosin heavy chain
44	89.5	18.1	559	1 S55383	peptidylprolyl iso
45	89.5	18.1	990	2 H88733	protein F32E10.3 [

ALIGNMENTS

RESULT 1

A97887  
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; F  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.,  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID:g1  
C:Genetics:  
A:Gene: pspA

Query Match	85.1%	Score	420.5	DB	2	Length	619
Best Local Similarity	90.0%	Pred. NO.	3.1e-23				
Matches	90	Conservative	2	Mismatches	7	Indels	1
Gaps	1						
QY	1	LKEIDSESDYAKGFRAPLOSKLDKAKLKLKLEELSDKIDELDAIAKLCVQLKDA	60				
DB	223	LKEIDSESDYAKGFRAPLOSKLDKAKLKLKLEELSDKIDELDAIAKLE-DOLKAA	281				
QY	61	EGNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDPE	100				
DB	282	EENNVEDYFKEGLEKTTAEKKAELKAEADLKKAVDPE	321				

RESULT 2

A41971  
surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:g153840; PIDN:AAA2701  
A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.



## RESULT 10

S73852  
hypothetical protein MG218 homolog F10\_orf1818 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S73852  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73852  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1818 <HIM>  
A:Cross-references: UNIPROT:P75471; EMBL:AB000051; GB:U00089; NID:gl674211; PIDN:AA89617  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
C:Genetic code: SGC3  
C:Superfamily: Mycoplasma genitalium hypothetical protein MG218

Query Match 20.0%; Score 99; DB 1; Length 1818;  
Best Local Similarity 27.0%; Pred. No. 14;  
Matches 31; Conservative 21; Mismatches 35; Indels 28; Gaps 4;

Qy 14 KEGFRAPLQSKLDKAKKAKL-----SKLESLDKIDELDAEIAKL--ECVQLKDAEGN-- 65  
Db 1364 KEGSLOGILQKLSUKTKQIEQFESKLQOREKLDQRQTTLSKHLRELKAQNEATAHKNRE 1423

Qy 66 ---VEAYFKEGKLETTAAK-----KAELEKABADLKKADEPE 100  
Db 1424 VLEIENYKELQRLITTEKSEFDNKNRLPEYFRKIRNEIEKKAHKTIVLEETQ 1478

RESULT 11  
AC1754  
capsid protein [bacteriophage bIL285] homolog lin2576 [imported] - Listeria innocua (str  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AC1754  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tisseret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A>Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1754  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-395 <GLA>  
A:Cross-references: UNIPROT:Q928F9; GB:AL592022; PIDN:CAC97803.1; PID:gl6415098; GSPDB:C  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin2576

Query Match 19.7%; Score 97.5; DB 2; Length 395;  
Best Local Similarity 30.7%; Pred. No. 4;  
Matches 27; Conservative 18; Mismatches 30; Indels 13; Gaps 2;

Qy 21 LQSKLDKAKKLSKLESLDKIDELDAEIAKLECVQLKDAEGN-- 71  
Db 8 IQKKLDSKRGLEELTRSKFKQGVLTTR----ALEEAKTEEVETVSKSDELETEK 63

Qy 72 EGGLEKTTAAKKALEKABADLKKADEP 99  
Db 64 EKLSKERDELTAKELEKELEANDKP 91

RESULT 12  
AS7013  
early endosome antigen 1 - human  
N:Alternate names: endosome-associated protein

C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: A57013; S44243  
R:Mu, F.T.; Callaghan, J.M.; Steele-Mortimer, O.; Stenmark, H.; Parton, R.G.; Campbell, I  
J. Biol. Chem. 270, 13503-13511, 1995  
A>Title: EEA1, an early endosome-associated protein. EEA1 is a conserved alpha-helical p  
A:Reference number: A57013; MUID:95286647; PMID:7768953  
A:Accession: A57013  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-1410 <RES>  
A:Cross-references: UNIPROT:Q15075; GB:L40157; NID:gi016367; PIDN:AAA79121.1; PID:gi01636  
R:Seelig, H.P.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S44243  
A:Accession: S44243  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-254, 'C', 256-257, 'LO', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 577-5  
A:Cross-references: EMBL:X78998; NID:g475933; PIDN:CAA55632.1; PID:g475934  
C:Genetics:  
A:Gene: GDB:EEA1  
A:Cross-references: GDB:I369996  
C:Superfamily: human early endosome antigen 1  
C:Keywords: calmodulin binding; endocytosis; metal binding; peripheral membrane protein;

Query Match 19.5%; Score 96.5; DB 1; Length 1410;  
Best Local Similarity 35.6%; Pred. No. 16;  
Matches 32; Conservative 16; Mismatches 31; Indels 11; Gaps 3;

Qy 9 SEDYAKGGRAPLQSKLD-AKAKLSKLESLDKIDELDAEIAKLECVQLKDAEGN-- 67  
Db 653 SAAAKTAQRADLQNLHDTAQNALQKHQLNKITTLQDQVTAQLQDKQ----EHCSQL 708

Qy 68 AYFKE-----GLEKTTAAKKALEKAEAD 91  
Db 709 SHLKEYKEYLSLEQTEELGGQIKKLEAD 738

RESULT 13  
A46173  
Mip4 protein - Streptococcus sp. (group A)  
C:Species: Streptococcus sp.  
C:Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 30-May-1997  
C:Accession: A46173  
R:O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.  
Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992  
A>Title: Two major classes in the M protein family in group A streptococci.  
A:Reference number: A46173; MUID:92409576; PMID:1528877  
A:Contents: group A  
A:Accession: A46173  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-388 <OIT>  
A:Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBIP:114064)  
C:Superfamily: M5 protein

Query Match 19.3%; Score 95.5; DB 2; Length 388;  
Best Local Similarity 28.1%; Pred. No. 5.4;  
Matches 43; Conservative 15; Mismatches 42; Indels 53; Gaps 4;

Qy 1 LKSIDSDSEYAKGFR-----PLQSKLDKAKKLSKLE----- 36  
Db 171 LKQODASKTEIETAKLQSEATLENLGSAKRELTQLAKLDTATAEKAKLESQVTTLENL 230

Qy 37 -----ELSDKIDELDAEIAKLECVQLKDAEGN--VEAYFKE-----GLEKT 77  
Db 231 LGSAKRELTDLQAKLDAANAEEKLSQQAATLEKQLEATKTELADLQAKLAATNQEKEL 290

Qy 78 TAEKKA-----ELEKAEADLKKADEPE 100  
Db 291 EAEAKALKEQAKQAEALAKLAKDASGAQKPD 323

```

RESULT 14
A33939
Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)
C:Species: Streptococcus sp.
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 26-Aug-1999
C:Accession: A33939
R:Heath, D.G.; Cleary, P.P.
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989
A:Title: Fc-receptor and M-protein genes of group A streptococci are products of gene du
A:Reference number: A33939; MUID:89282846; PMID:2660147
A:Accession: A33939
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <HEA>
A:Cross-references: GB:M22532; NID:G153628; PIDN:AAB95296.1; PID:G552003
C:Superfamily: M5 protein
C:Keywords: immunoglobulin receptor

Query Match 19.3%; Score 95.5; DB 2; Length 405;
Best Local Similarity 28.1%; Pred. NO. 5.7;
Matches 43; Conservative 15; Mismatches 42; Indels 53; Gaps 4;

QY 1 LKEIDESDSEYAKGFR-----PLOSKLDAKAKLSKLE----- 36
DB 208 LKQDASKTEETIAKLQSEATLENLGSARKRLTDLOAKLDTATAEKAKLESQVTTLENL 267
QY 37 -----ELSDKIDELDAEIAKLCVQLKDAEGNNVNEAYFKB-----GLEKT 77
DB 268 LGSARKRLTDLOAKLDANNAEKEQLSQAALEKLEATKKEKLADLOAKLAATNOEKL 327
QY 78 TAEKKA-----ELEKAEADLKKAVIDPE 100
DB 328 EAEAKALKQLAKQAEELAKLKADKASGAQKPD 360

RESULT 15
T05409
hypothetical protein F10M6.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05409
R:Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15414
A:Accession: T05409
A:Molecule type: DNA
A:Residues: 1-764 <BEV>
A:Cross-references: UNIPROT:O49371; EMBL:AL021811
A:Experimental source: cultivar Columbia; BAC clone F10M6
C:Genetics:
A:Map position: 4
A>Note: F10M6.170

Query Match 19.3%; Score 95.5; DB 2; Length 764;
Best Local Similarity 30.7%; Pred. NO. 11;
Matches 35; Conservative 22; Mismatches 40; Indels 17; Gaps 4;

QY 2 KEIDESDSEYAKGFRPLOSKLDAKAKLSKL-EELSDKIDELDAEIAKLCVQLKDA 60
DB 163 REIEELKHKLRRDEREALQSSLTKBELEKMRQEIANKRSKEYSMATSEFESKSLLS 222
QY 61 EGNNV-----EAYF-----KEGLEKTTAEKKALEK---AAADLKAVDE 98
DB 223 KANEVVVKQGEIYALQRALEKEBELEISKATKKLEKLEKREANLUKQTEE 276

Search completed: June 18, 2005, 17:03:57
Job time : 14.113 secs

```

**This Page Blank (uspto)**



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	459.5	93.0	739	2	Q9RQT4	streptococc
2	459.5	93.0	820	2	Q9RQT1	streptococc
3	459.5	93.0	929	2	Q9KK19	streptococc
4	459.5	93.0	929	2	Q9ZAY5	streptococc
5	450.5	91.2	437	2	Q9LAY4	streptococc
6	444.5	90.0	395	2	Q9LAY2	streptococc
7	444.5	90.0	408	2	Q9LAY0	streptococc
8	440.5	89.2	249	2	Q9L575	streptococc
9	433.5	87.8	224	2	Q8GNS8	streptococc
10	430.5	87.1	99	2	Q8QXK4	streptococc
11	427.5	86.5	426	2	Q9LAY5	streptococc
12	420.5	85.1	619	2	Q54972	streptococc
13	420.5	85.1	619	2	Q8DR10	streptococc
14	419	84.8	869	2	Q9KK27	streptococc
15	412.5	83.5	417	2	Q9LAY3	streptococc
16	398.5	80.7	415	2	Q9LAY1	streptococc
17	335	67.8	394	2	Q9LAY6	streptococc
18	335	67.8	395	2	Q9LAZ1	streptococc
19	332	67.2	225	2	Q9L591	streptococc
20	332	67.2	246	2	Q9L578	streptococc
21	327	66.2	255	2	Q9L581	streptococc
22	327	66.2	255	2	Q9L5B6	streptococc
23	326	66.0	194	2	Q9L5B5	streptococc
24	326	66.0	218	2	Q6UEB2	streptococc
25	326	66.0	222	2	Q9L577	streptococc
26	326	66.0	233	2	Q9L568	streptococc
27	326	66.0	236	2	Q9L569	streptococc
28	326	66.0	243	2	Q9L564	streptococc
29	326	66.0	243	2	Q9L567	streptococc
30	326	66.0	244	2	Q9L565	streptococc
31	326	66.0	247	2	Q9L566	streptococc



DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E134;  
RX MEDLINE=20448953; PubMed=10992499;  
R DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae."  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071811; AAF27707.1; -;  
DR InterPro; IPR002479; CW\_binding.  
DR Pfam; PF01473; CW\_binding\_1; 1.  
FT NON\_TER 437 437  
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;  
  
Query Match 91.2%; Score 450.5; DB 2; Length 437;  
Best Local Similarity 94.0%; Pred. No. 8.6e-22; Indels 1; Gaps 1;  
Matches 94; Conservative 0; Mismatches 5;  
  
QY 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
Db 235 LKEIDESSEDYKXKEGLRAPLQSKLDTKAKLSKLEELSDKIDELDAETAK-HVOLKDA 293  
  
QY 61 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 100  
Db 294 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 333  
  
RESULT 6  
Q9LAY2 PRELIMINARY; PRT; 395 AA.  
AC Q9LAY2;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EF6796;  
RX MEDLINE=20448953; PubMed=10992499;  
R DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae."  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071813; AAF27709.1; -;  
DR InterPro; IPR009053; Prefoldin.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 395 395  
SQ SEQUENCE 395 AA; 42963 MW; 58B6EF956BCBCC1E CRC64;  
  
Query Match 90.0%; Score 444.5; DB 2; Length 395;  
Best Local Similarity 93.0%; Pred. No. 1.9e-21;  
Matches 93; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
  
QY 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
Db 225 LKEINESSEDYAKGFRAPLQSKLDKAKKLSKLEELSGKIEELDAETAELE-VOLKDA 283

QY 61 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 100  
Db 284 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 323  
  
RESULT 7  
Q9LAY0 PRELIMINARY; PRT; 408 AA.  
AC Q9LAY0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BG9163;  
RX MEDLINE=20448953; PubMed=10992499;  
R DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae."  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071815; AAF27711.1; -;  
DR InterPro; IPR009053; Prefoldin.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
FT NON\_TER 408 408  
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;  
  
Query Match 90.0%; Score 444.5; DB 2; Length 408;  
Best Local Similarity 93.0%; Pred. No. 2e-21; Indels 1; Gaps 1;  
Matches 93; Conservative 4; Mismatches 2;  
  
QY 1 LKEIDESSEDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLECVOLKDA 60  
Db 228 LKEINESSEDYAKGFRAPLQSKLDKAKKLSKLEELSGKIEELDAETAELE-VOLKDA 286  
  
QY 61 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 100  
Db 287 EGNNVVEAYFKGEGLEKTTAAKAELEKAEADLKKAADLKKAVDEPE 326  
  
RESULT 8  
Q9L575 PRELIMINARY; PRT; 249 AA.  
AC Q9L575;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=195;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
pneumococcal strains in the United States and of internationally  
disseminated clones."  
RL J. Clin. Microbiol. 38:3663-3669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=195;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.

```

DR EMBL: AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match      89.2%; Score 440.5; DB 2; Length 249;
Best Local Similarity 92.0%; Pred. No. 2.3e-21;
Matches 92; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYAKGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 LKEIDSDSDYAKGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAEIE-VQLKDA 132

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 133 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 172

RESULT 9
Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicunzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
   Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match      87.8%; Score 433.5; DB 2; Length 224;
Best Local Similarity 90.0%; Pred. No. 5.9e-21;
Matches 90; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYAKGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17 LKIDNESDSDYKVGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAEIE-VQLKDA 75

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 76 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 115

RESULT 10
Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;

```

---

```

RN EMBL: AF255552; AAF68105.1; -.
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
   immunization with DNA vaccines against Streptococcus pneumoniae
   expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match      87.1%; Score 430.5; DB 2; Length 99;
Best Local Similarity 90.0%; Pred. No. 4.4e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYAKGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSDSDYKVGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAEIE-VQLKDA 59

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 60 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVNEPE 99

RESULT 11
Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
   in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_ald_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match      86.5%; Score 427.5; DB 2; Length 426;
Best Local Similarity 89.0%; Pred. No. 2.6e-20;
Matches 89; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYAKGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 215 LKIDNESDSDYKVGFRAPLQSKLDAAKAKLSKLELSKIDELDAEIAEIE-VQLKDA 273

Qy 61 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 274 EGNNVVAYFKGLEKTTAEKKAELKAEADLKAVDEPE 313

RESULT 12

```

Q54972  
 ID Q54972; PRELIMINARY; PRT; 619 AA.  
 AC Q54972;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Pneuomococcal surface protein A precursor.  
 GN Name=pspA;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92105030; PubMed=1729249;  
 RA Yother J., Briles D.E.;  
 RT "Structural properties and evolutionary relationships of PspA, a  
 RT surface protein of Streptococcus pneumoniae, as revealed by sequence  
 RT analysis.";  
 RL J. Bacteriol. 174:601-609(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yother J., Briles D.E.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M74122; AAA27018.1; -;  
 DR PIR; A41971; A41971.  
 DR PIR; A97887; A97887.  
 DR HSSP; P06653; IHXC.  
 DR InterPro; IPR002479; CW binding.  
 DR InterPro; IPR002345; Lipocalin.  
 DR Pfam; PF01473; CW binding\_1; 10.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_3.  
 DR SIGNAL.  
 KW Signal.  
 FT SIGNAL.  
 FT CHAIN 1 31 Potential.  
 FT CHAIN 32 619 pneumococcal surface protein A.  
 SQ SEQUENCE 619 AA; 68605 MW; 5A8BDB40C2841CA CRC64;  
 Query Match 85.1%; Score 420.5; DB 2; Length 619;  
 Best Local Similarity 90.0%; Pred. No. 1.1e-19;  
 Matches 90; Conservative 2; Mismatches 7; Indels 1; Gaps 1;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLECVOLKDA 60  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLE- DQKAA 281  
 QY 61 EGNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 100  
 DB 282 EENNVEDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 321  
 Query Match 85.1%; Score 420.5; DB 2; Length 619;  
 Best Local Similarity 90.0%; Pred. No. 1.1e-19;  
 Matches 90; Conservative 2; Mismatches 7; Indels 1; Gaps 1;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLECVOLKDA 60  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLE- DQKAA 281  
 QY 61 EGNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 100  
 DB 282 EENNVEDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 321  
 RESULT 13  
 Q8DR10  
 ID Q8DR10 PRELIMINARY; PRT; 619 AA.  
 AC Q8DR10;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Surface protein pspA  
 GN Name=pspA; OrderedlocusNames=sp0121;  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21429245; PubMed=11544234;  
 RX DOI=10.1128/JB.183.19.5709-5717.2001;  
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Bargett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McAdams S.N., McHenry M., McEaster K., Mundy C.W., Niclas T.I.,  
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,

RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,  
 RA Glass J.I.;  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 DR EMBL; AE008396; AAK98925.1; -;  
 DR PIR; A41971; A41971.  
 DR PIR; A97887; A97887.  
 DR HSSP; P06653; IHXC.  
 DR InterPro; IPR002479; CW binding.  
 DR InterPro; IPR002345; Lipocalin.  
 DR Pfam; PF01473; CW binding\_1; 10.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_3.  
 KW Complete proteome.  
 SQ SEQUENCE 619 AA; 68605 MW; 5A8BDB40C2841CA CRC64;  
 Query Match 85.1%; Score 420.5; DB 2; Length 619;  
 Best Local Similarity 90.0%; Pred. No. 1.1e-19;  
 Matches 90; Conservative 2; Mismatches 7; Indels 1; Gaps 1;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLECVOLKDA 60  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLE- DQKAA 281  
 QY 61 EGNNNVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 100  
 DB 282 EENNVEDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 321  
 RESULT 14  
 Q9KK27  
 ID Q9KK27 PRELIMINARY; PRT; 869 AA.  
 AC Q9KK27;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Surface protein PspC.  
 GN Name=pspC;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=95;  
 RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;  
 RA Iannelli F., Oggioni M.R., Pozzi G.;  
 RT "Allelic variation in the highly polymorphic locus pspC of  
 RT Streptococcus pneumoniae.";  
 RL Gene 284:63-71(2002).  
 DR EMBL; AF154032; AAF73801.1; -;  
 DR HSSP; P06653; IHXC.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR002479; CW binding.  
 DR InterPro; IPR005877; Gpos YSIRK.  
 DR InterPro; IPR007756; RICH.  
 DR Pfam; PF01473; CW binding\_1; 8.  
 DR Pfam; PF05062; RICH; 2.  
 DR Pfam; PF04650; YSIRK\_signal; 1.  
 DR TIGRFAMs; TIGR01168; YSIRK\_signal; 1.  
 SQ SEQUENCE 869 AA; 98732 MW; AFP2B504347E0220 CRC64;  
 Query Match 84.8%; Score 419; DB 2; Length 869;  
 Best Local Similarity 86.4%; Pred. No. 1.8e-19;  
 Matches 89; Conservative 1; Mismatches 5; Indels 8; Gaps 2;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLECVOLKDA 57  
 DB 537 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKEELSDKIDELDAETAKLE- VNCNLRSQL 591  
 QY 58 KDAEGNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 100  
 DB 592 KDAEGNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 634

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PpA (Fragment).
GN Name=ppa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBL_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
EX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PpA: mosaic genes and evidence for past recombination
RL in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
FT NON_TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match      83.5%; Score 412.5; DB 2; Length 417;
Best Local Similarity 88.0%; Pred. No. 2.4e-19;
Matches 88; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Qy      1 LKEIDSDSDYAKGFPAPLQSKLDAKAKLSKLELSDKIDELDAEIAKLECVQLKDA 60
Db      213 LKEIDSDSDYVKEGFPAPLQSKLDAKAKLSKLELSDKIDELDAEIAKLE-DQLKAA 271

Qy      61 EGNNVVAYPKGKLEKTTAEKKAELEKAEADLKKAVDEPE 100
Db      272 EENNVEDYFKGLEKTIKAEKAELEKTEADLKKAVNEPE 311
```

Search completed: June 18, 2005, 17:01:37  
Job time : 61.961 secs